

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> :  C07H 21/04, C07K 1/00, 14/00, C12N 1/21, 15/00, 15/09, 15/63, 15/70, C12P 19/34		A1	(11) International Publication Number: <b>WO 00/52027</b>  (43) International Publication Date: 8 September 2000 (08.09.00)		
(21) International Application Number: PCT/US00/05432		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).			
(22) International Filing Date: 2 March 2000 (02.03.00)					
(30) Priority Data: 60/122,389 2 March 1999 (02.03.99) US 60/126,049 23 March 1999 (23.03.99) US 60/136,744 28 May 1999 (28.05.99) US					
(71) Applicant: LIFE TECHNOLOGIES, INC. [US/US]; 9800 Medical Center Drive, Rockville, MD 20850 (US).					
(72) Inventors: HARTLEY, James, L.; 7409 Hillside Drive, Frederick, MD 21702 (US). BRASCH, Michael, A.; 20931 Sunnymacres Road, Gaithersburg, MD 20882 (US). TEMPLE, Gary, F.; 114 Ridge Road, Washington Grove, MD 20882 (US). CHEO, David; 2006 Baltimore Road, #21, Rockville, MD 20851 (US).		Published <i>With international search report.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>			
(74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).					
(54) Title: COMPOSITIONS AND METHODS FOR USE IN RECOMBINATIONAL CLONING OF NUCLEIC ACIDS					
(57) Abstract					
<p>The present invention relates generally to compositions and methods for use in recombinational cloning of nucleic acid molecules. In particular, the invention relates to nucleic acid molecules encoding one or more recombination sites or portions thereof, to nucleic acid molecules comprising one or more of these recombination site nucleotide sequences and optionally comprising one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides using the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof. The invention also relates to the use of these compositions in methods for recombinational cloning of nucleic acids, <i>in vitro</i> and <i>in vivo</i>, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.</p>					

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

# Compositions and Methods for Use in Recombinational Cloning of Nucleic Acids

5

## BACKGROUND OF THE INVENTION

### *Field of the Invention*

The present invention relates generally to recombinant DNA technology.

More particularly, the present invention relates to compositions and methods for use in recombinational cloning of nucleic acid molecules. The invention relates specifically to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides and RNAs encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments. More particularly, the antibodies of the invention may be used to identify and/or purify proteins or fusion proteins encoded by the nucleic acid molecules or vectors of the invention, or to identify and/or purify the nucleic acid molecules of the invention.

**Related Art**

5                   **Site-specific recombinases.** Site-specific recombinases are proteins that are present in many organisms (e.g. viruses and bacteria) and have been characterized to have both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, e.g., Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).

10

Numerous recombination systems from various organisms have been described. See, e.g., Hoess *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Abremski *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Campbell, J. *Bacteriol.* 174(23):7495 (1992); Qian *et al.*, *J. Biol. Chem.* 267(11):7794 (1992);  
15 Araki *et al.*, *J. Mol. Biol.* 225(1):25 (1992); Maeser and Kahnmann *Mol. Gen. Genet.* 230:170-176 (1991); Esposito *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997).

20                  Many of these belong to the integrase family of recombinases (Argos *et al.* *EMBO J.* 5:433-440 (1986); Voznyanov *et al.*, *Nucl. Acids Res.* 27:930 (1999)).

Perhaps the best studied of these are the Integrase/att system from bacteriophage  $\lambda$  (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/loxP system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the Saccharomyces cerevisiae 2  $\mu$  circle plasmid (Broach *et al.* *Cell* 29:227-234 (1982)).

25

Backman (U.S. Patent No. 4,673,640) discloses the *in vivo* use of  $\lambda$  recombinase to recombine a protein producing DNA segment by enzymatic site-specific recombination using wild-type recombination sites attB and attP.

Hasan and Szybalski (*Gene* 56:145-151 (1987)) discloses the use of  $\lambda$  Int recombinase *in vivo* for intramolecular recombination between wild type attP and attB sites which flank a promoter. Because the orientations of these sites are

inverted relative to each other, this causes an irreversible flipping of the promoter region relative to the gene of interest.

5 Palazzolo *et al.* *Gene* 88:25-36 (1990), discloses phage lambda vectors having bacteriophage  $\lambda$  arms that contain restriction sites positioned outside a cloned DNA sequence and between wild-type *loxP* sites. Infection of *E. coli* cells that express the Cre recombinase with these phage vectors results in recombination between the *loxP* sites and the *in vivo* excision of the plasmid replicon, including the cloned cDNA.

10 Pósfai *et al.* (*Nucl. Acids Res.* 22:2392-2398 (1994)) discloses a method for inserting into genomic DNA partial expression vectors having a selectable marker, flanked by two wild-type FRT recognition sequences. FLP site-specific recombinase as present in the cells is used to integrate the vectors into the genome at predetermined sites. Under conditions where the replicon is functional, this cloned genomic DNA can be amplified.

15 Bebee *et al.* (U.S. Patent No. 5,434,066) discloses the use of site-specific recombinases such as Cre for DNA containing two *loxP* sites for *in vivo* recombination between the sites.

20 Boyd (*Nucl. Acids Res.* 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intermolecular ligation to a dephosphorylated vector that contains a wild-type *loxP* site acted upon by a Cre site-specific recombinase present in *E. coli* host cells.

25 Waterhouse *et al.* (WO 93/19172 and *Nucleic Acids Res.* 21 (9):2265 (1993)) disclose an *in vivo* method where light and heavy chains of a particular antibody were cloned in different phage vectors between *loxP* and *loxP 511* sites and used to transfect new *E. coli* cells. Cre, acting in the host cells on the two parental molecules (one plasmid, one phage), produced four products in equilibrium: two different cointegrates (produced by recombination at either *loxP* or *loxP 511* sites), and two daughter molecules, one of which was the desired product.

30 Schlake & Bode (*Biochemistry* 33:12746-12751 (1994)) discloses an *in vivo* method to exchange expression cassettes at defined chromosomal locations, each flanked by a wild type and a spacer-mutated FRT recombination site. A

double-reciprocal crossover was mediated in cultured mammalian cells by using this FLP/FRT system for site-specific recombination.

Hartley *et al.* (U.S. Patent No. 5,888,732) disclose compositions and methods for recombinational exchange of nucleic acid segments and molecules, including for use in recombinational cloning of a variety of nucleic acid molecules *in vitro* and *in vivo*, using a combination of wildtype and mutated recombination sites and recombination proteins.

**Transposases.** The family of enzymes, the transposases, has also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of transposons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the *in vivo* movement of DNA segments between replicons (Lucklow *et al.*, *J. Virol.* 67:4566-4579 (1993)).

Devine and Boeke *Nucl. Acids Res.* 22:3765-3772 (1994), discloses the construction of artificial transposons for the insertion of DNA segments, *in vitro*, into recipient DNA molecules. The system makes use of the integrase of yeast TY1 virus-like particles. The DNA segment of interest is cloned, using standard methods, between the ends of the transposon-like element TY1. In the presence of the TY1 integrase, the resulting element integrates randomly into a second target DNA molecule.

**Recombination Sites.** Also key to the integration/recombination reactions mediated by the above-noted recombination proteins and/or transposases are recognition sequences, often termed "recombination sites," on the DNA molecules participating in the integration/recombination reactions. These recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by the recombination proteins during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.*

5 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombination protein  $\lambda$  Int. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region, while *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993); see also U.S. Patent No. 5,888,732, which is incorporated by reference herein.

10 **DNA cloning.** The cloning of DNA segments currently occurs as a daily routine in many research labs and as a prerequisite step in many genetic analyses. The purpose of these clonings is various, however, two general purposes can be considered: (1) the initial cloning of DNA from large DNA or RNA segments (chromosomes, YACs, PCR fragments, mRNA, etc.), done in a relative handful of known vectors such as pUC, pGem, pBlueScript, and (2) the subcloning of these DNA segments into specialized vectors for functional analysis. A great deal of time and effort is expended both in the transfer of DNA segments from the initial cloning vectors to the more specialized vectors. This transfer is called subcloning.

15

20 The basic methods for cloning have been known for many years and have changed little during that time. A typical cloning protocol is as follows:

25

- (1) digest the DNA of interest with one or two restriction enzymes;
- (2) gel purify the DNA segment of interest when known;
- (3) prepare the vector by cutting with appropriate restriction enzymes, treating with alkaline phosphatase, gel purify etc., as appropriate;
- (4) ligate the DNA segment to the vector, with appropriate controls to eliminate background of uncut and self-ligated vector;
- (5) introduce the resulting vector into an *E. coli* host cell;
- (6) pick selected colonies and grow small cultures overnight;
- (7) make DNA minipreps; and

(8) analyze the isolated plasmid on agarose gels (often after diagnostic restriction enzyme digestions) or by PCR.

The specialized vectors used for subcloning DNA segments are functionally diverse. These include but are not limited to: vectors for expressing nucleic acid molecules in various organisms; for regulating nucleic acid molecule expression; for providing tags to aid in protein purification or to allow tracking of proteins in cells; for modifying the cloned DNA segment (*e.g.*, generating deletions); for the synthesis of probes (*e.g.*, riboprobes); for the preparation of templates for DNA sequencing; for the identification of protein coding regions; for the fusion of various protein-coding regions; to provide large amounts of the DNA of interest, *etc.* It is common that a particular investigation will involve subcloning the DNA segment of interest into several different specialized vectors.

As known in the art, simple subclonings can be done in one day (*e.g.*, the DNA segment is not large and the restriction sites are compatible with those of the subcloning vector). However, many other subclonings can take several weeks, especially those involving unknown sequences, long fragments, toxic genes, unsuitable placement of restriction sites, high backgrounds, impure enzymes, *etc.* Subcloning DNA fragments is thus often viewed as a chore to be done as few times as possible.

Several methods for facilitating the cloning of DNA segments have been described, *e.g.*, as in the following references.

Ferguson, J., *et al.* *Gene* 16:191 (1981), discloses a family of vectors for subcloning fragments of yeast DNA. The vectors encode kanamycin resistance. Clones of longer yeast DNA segments can be partially digested and ligated into the subcloning vectors. If the original cloning vector conveys resistance to ampicillin, no purification is necessary prior to transformation, since the selection will be for kanamycin.

Hashimoto-Gotoh, T., *et al.* *Gene* 41:125 (1986), discloses a subcloning vector with unique cloning sites within a streptomycin sensitivity gene; in a streptomycin-resistant host, only plasmids with inserts or deletions in the dominant sensitivity gene will survive streptomycin selection.

Accordingly, traditional subcloning methods, using restriction enzymes and ligase, are time consuming and relatively unreliable. Considerable labor is expended, and if two or more days later the desired subclone can not be found among the candidate plasmids, the entire process must then be repeated with alternative conditions attempted. Although site specific recombinases have been used to recombine DNA *in vivo*, the successful use of such enzymes *in vitro* was expected to suffer from several problems. For example, the site specificities and efficiencies were expected to differ *in vitro*; topologically linked products were expected; and the topology of the DNA substrates and recombination proteins was expected to differ significantly *in vitro* (see, e.g., Adams *et al*, *J. Mol. Biol.* 226:661-73 (1992)). Reactions that could go on for many hours *in vivo* were expected to occur in significantly less time *in vitro* before the enzymes became inactive. In addition, the stabilities of the recombination enzymes after incubation for extended periods of time in *in vitro* reactions was unknown, as were the effects of the topologies (*i.e.*, linear, coiled, supercoiled, etc.) of the nucleic acid molecules involved in the reaction. Multiple DNA recombination products were expected in the biological host used, resulting in unsatisfactory reliability, specificity or efficiency of subcloning. Thus, *in vitro* recombination reactions were not expected to be sufficiently efficient to yield the desired levels of product.

Accordingly, there is a long felt need to provide an alternative subcloning system that provides advantages over the known use of restriction enzymes and ligases.

25

## SUMMARY OF THE INVENTION

30

The present invention relates to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules comprising one or more of the recombination site nucleotide sequences or portions thereof and one or more additional physical or functional nucleotide sequences, such as those

encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (e.g., one or more promoters, enhancers, or repressors), one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (e.g., GST, His<sub>6</sub> or thioredoxin), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more desired proteins or peptides encoded by a gene or a portion of a gene, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences.

The invention also relates to primer nucleic acid molecules comprising the recombination site nucleotide sequences of the invention (or portions thereof), and to such primer nucleic acid molecules linked to one or more target-specific (e.g., one or more gene-specific) primer nucleic acid sequences. Such primers may also comprise sequences complementary or homologous to DNA or RNA sequences to be amplified, e.g., by PCR, RT-PCR, etc. Such primers may also comprise sequences or portions of sequences useful in the expression of protein genes (ribosome binding sites, localization signals, protease cleavage sites, repressor binding sites, promoters, transcription stops, stop codons, etc.). Said primers may also comprise sequences or portions of sequences useful in the manipulation of DNA molecules (restriction sites, transposition sites, sequencing primers, etc.). The primers of the invention may be used in nucleic acid synthesis and preferably are used for amplification (e.g., PCR) of nucleic acid molecules. When the primers of the invention include target- or gene-specific sequences (any sequence contained within the target to be synthesized or amplified including translation signals, gene sequences, stop codons, transcriptional signals (e.g., promoters) and the like), amplification or synthesis of target sequences or genes may be accomplished. Thus, the invention relates to synthesis of a nucleic acid molecules comprising mixing one or more primers of the invention with a nucleic acid

-9-

template, and incubating said mixture under conditions sufficient to make a first nucleic acid molecule complementary to all or a portion of said template. Thus, the invention relates specifically to a method of synthesizing a nucleic acid molecule comprising:

5                     (a) mixing a nucleic acid template with a polypeptide having polymerase activity and one or more primers comprising one or more recombination sites or portions thereof; and

10                   (b) incubating said mixture under conditions sufficient to synthesize a first nucleic acid molecule complementary to all or a portion of said template and which preferably comprises one or more recombination sites or portions thereof.

15                  Such method of the invention may further comprise incubating said first synthesized nucleic acid molecule under conditions sufficient to synthesize a second nucleic acid molecule complementary to all or a portion of said first nucleic acid molecule. Such synthesis may provide for a first nucleic acid molecule having a recombination site or portion thereof at one or both of its termini.

20                  In a preferred aspect, for the synthesis of the nucleic acid molecules, at least two primers are used wherein each primer comprises a homologous sequence at its terminus and/or within internal sequences of each primer (which may have a homology length of about 2 to about 500 bases, preferably about 3 to about 100 bases, about 4 to about 50 bases, about 5 to about 25 bases and most preferably about 6 to about 18 base overlap). In a preferred aspect, the first such primer comprises at least one target-specific sequence and at least one recombination site or portion thereof while the second primer comprises at least one recombination site or portion thereof. Preferably, the homologous regions between the first and second primers comprise at least a portion of the recombination site. In another aspect, the homologous regions between the first and second primers may comprise one or more additional sequences, e.g., expression signals, translational start motifs, or other sequences adding functionality to the desired nucleic acid sequence upon amplification. In practice, two pairs of primers prime synthesis or amplification of a nucleic acid molecule. In a preferred aspect, all or at least a portion of the synthesized or amplified nucleic acid molecule will be homologous

25

30

to all or a portion of the template and further comprises a recombination site or a portion thereof at at least one terminus and preferably both termini of the synthesized or amplified molecule. Such synthesized or amplified nucleic acid molecule may be double stranded or single stranded and may be used in the recombinational cloning methods of the invention. The homologous primers of the invention provide a substantial advantage in that one set of the primers may be standardized for any synthesis or amplification reaction. That is, the primers providing the recombination site sequences (without the target specific sequences) can be pre-made and readily available for use. This in practice allows the use of shorter custom made primers that contain the target specific sequence needed to synthesize or amplify the desired nucleic acid molecule. Thus, this provides reduced time and cost in preparing target specific primers (e.g., shorter primers containing the target specific sequences can be prepared and used in synthesis reactions). The standardized primers, on the other hand, may be produced in mass to reduce cost and can be readily provided (e.g., in kits or as a product) to facilitate synthesis of the desired nucleic acid molecules.

Thus, in one preferred aspect, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

More specifically, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

In a more preferred aspect, the invention relates to a method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and one or more first primers comprising at least a portion of a recombination site and a template specific sequence (complementary to or capable of hybridizing to said template);
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one and preferably both termini of said molecules;
- (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or

complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and

(d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one and preferably both termini of said molecules.

5

10

15

20

25

30

The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, primers, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.

The antibodies of the invention may have particular use to identify and/or purify peptides or proteins (including fusion proteins produced by the invention), and to identify and/or purify the nucleic acid molecules of the invention or portions thereof.

The methods for *in vitro* or *in vivo* recombinational cloning of nucleic acid molecule generally relate to recombination between at least a first nucleic acid molecule having at least one recombination site and a second nucleic acid molecule having at least one recombination site to provide a chimeric nucleic acid molecule. In one aspect, the methods relate to recombination between a first vector having at least one recombination site and a second vector having at least one recombination site to provide a chimeric vector. In another aspect, a nucleic acid molecule having at least one recombination site is combined with a vector having at least one recombination site to provide a chimeric vector. In a most preferred aspect, the nucleic acid molecules or vectors used in recombination

comprise two or more recombination sites. In a more specific embodiment of the invention, the recombination methods relate to a Destination Reaction (also referred to herein as an "LR reaction") in which recombination occurs between an Entry clone and a Destination Vector. Such a reaction transfers the nucleic acid molecule of interest from the Entry Clone into the Destination Vector to create an Expression Clone. The methods of the invention also specifically relate to an Entry or Gateward reaction (also referred to herein as a "BP reaction") in which an Expression Clone is recombined with a Donor vector to produce an Entry clone. In other aspects, the invention relates to methods to prepare Entry clones by combining an Entry vector with at least one nucleic acid molecule (e.g., gene or portion of a gene). The invention also relates to conversion of a desired vector into a Destination Vector by including one or more (preferably at least two) recombination sites in the vector of interest. In a more preferred aspect, a nucleic acid molecule (e.g., a cassette) having at least two recombination sites flanking a selectable marker (e.g., a toxic gene or a genetic element preventing the survival of a host cell containing that gene or element, and/or preventing replication, partition or heritability of a nucleic acid molecule (e.g., a vector or plasmid) comprising that gene or element) is added to the vector to make a Destination Vector of the invention.

Preferred vectors for use in the invention include prokaryotic vectors, eukaryotic vectors, or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred prokaryotic vectors for use in the invention include but are not limited to vectors which may propagate and/or replicate in gram negative and/or gram positive bacteria, including bacteria of the genera *Escherichia*, *Salmonella*, *Proteus*, *Clostridium*, *Klebsiella*, *Bacillus*, *Streptomyces*, and *Pseudomonas* and preferably in the species *E. coli*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate in yeast cells, plant cells, mammalian cells, (particularly human and mouse), fungal cells, insect cells, nematode cells, fish cells and the like. Particular vectors of interest include but are not limited to cloning vectors, sequencing vectors, expression vectors, fusion vectors, two-hybrid vectors, gene therapy vectors, phage display vectors, gene-targeting vectors, PACs, BACs, YACs, MACs, and

reverse two-hybrid vectors. Such vectors may be used in prokaryotic and/or eukaryotic systems depending on the particular vector.

In another aspect, the invention relates to kits which may be used in carrying out the methods of the invention, and more specifically relates to cloning or subcloning kits and kits for carrying out the LR Reaction (e.g., making an Expression Clone), for carrying out the BP Reaction (e.g., making an Entry Clone), and for making Entry Clone and Destination Vector molecules of the invention. Such kits may comprise a carrier or receptacle being compartmentalized to receive and hold therein any number of containers. Such containers may contain any number of components for carrying out the methods of the invention or combinations of such components. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins or auxiliary factors or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix), one or more reaction buffers, one or more nucleotides, one or more primers of the invention, one or more restriction enzymes, one or more ligases, one or more polypeptides having polymerase activity (e.g., one or more reverse transcriptases or DNA polymerases), one or more proteinases (e.g., proteinase K or other proteinases), one or more Destination Vector molecules, one or more Entry Clone molecules, one or more host cells (e.g. competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3.1 host cells, such as *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells), instructions for using the kits of the invention (e.g., to carry out the methods of the invention), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, particularly one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites or portions thereof of the invention. Preferably, such nucleic acid molecules comprise at least two recombination sites which flank a selectable

marker (*e.g.*, a toxic gene and/or antibiotic resistance gene). In a preferred aspect, such nucleic acid molecules are in the form of a cassette (*e.g.*, a linear nucleic acid molecule comprising one or more and preferably two or more recombination sites or portions thereof).

5 Kits for inserting or adding recombination sites to nucleic acid molecules of interest may comprise one or more nucleases (preferably restriction endonucleases), one or more ligases, one or more topoisomerases, one or more polymerases, and one or more nucleic acid molecules or adapters comprising one or more recombination sites. Kits for integrating recombination sites into one or  
10 more nucleic acid molecules of interest may comprise one or more components (or combinations thereof) selected from the group consisting of one or more integration sequences comprising one or more recombination sites. Such integration sequences may comprise one or more transposons, integrating viruses, homologous recombination sequences, RNA molecules, one or more host cells  
15 and the like.

Kits for making the Entry Clone molecules of the invention may comprise any or a number of components and the composition of such kits may vary depending on the specific method involved. Such methods may involve inserting the nucleic acid molecules of interest into an Entry or Donor Vector by the recombinational cloning methods of the invention, or using conventional molecular biology techniques (*e.g.*, restriction enzyme digestion and ligation). In a preferred aspect, the Entry Clone is made using nucleic acid amplification or synthesis products. Kits for synthesizing Entry Clone molecules from amplification or synthesis products may comprise one or more components (or combinations thereof) selected from the group consisting of one or more Donor Vectors (*e.g.*, one or more attP vectors including, but not limited to, pDONR201 (Figure 49), pDONR202 (Figure 50), pDONR203 (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (Figure 53), and the like), one or more polypeptides having polymerase activity (preferably DNA polymerases and most preferably thermostable DNA polymerases), one or more proteinases, one or more reaction buffers, one or more nucleotides, one or more primers comprising one or  
20  
25  
30

more recombination sites or portions thereof, and instructions for making one or more Entry Clones.

Kits for making the Destination vectors of the invention may comprise any number of components and the compositions of such kits may vary depending on the specific method involved. Such methods may include the recombination methods of the invention or conventional molecular biology techniques (e.g., restriction endonuclease digestion and ligation). In a preferred aspect, the Destination vector is made by inserting a nucleic acid molecule comprising at least one recombination site (or portion thereof) of the invention (preferably a nucleic acid molecule comprising at least two recombination sites or portions thereof flanking a selectable marker) into a desired vector to convert the desired vector into a Destination vector of the invention. Such kits may comprise at least one component (or combinations thereof) selected from the group consisting of one or more restriction endonucleases, one or more ligases, one or more polymerases, one or more nucleotides, reaction buffers, one or more nucleic acid molecules comprising at least one recombination site or portion thereof (preferably at least one nucleic acid molecule comprising at least two recombination sites flanking at least one selectable marker, such as a cassette comprising at least one selectable marker such as antibiotic resistance genes and/or toxic genes), and instructions for making such Destination vectors.

The invention also relates to kits for using the antibodies of the invention in identification and/or isolation of peptides and proteins (which may be fusion proteins) produced by the nucleic acid molecules of the invention, and for identification and/or isolation of the nucleic acid molecules of the invention or portions thereof. Such kits may comprise one or more components (or combination thereof) selected from the group consisting of one or more antibodies of the invention, one or more detectable labels, one or more solid supports and the like.

Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of what is known in the art, in light of the following drawings and description of the invention, and in light of the claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts one general method of the present invention, wherein the starting (parent) DNA molecules can be circular or linear. The goal is to exchange the new subcloning vector D for the original cloning vector B. It is desirable in one embodiment to select for AD and against all the other molecules, including the Cointegrate. The square and circle are sites of recombination: e.g., *lox* (such as *loxP*) sites, *att* sites, etc. For example, segment D can contain expression signals, protein fusion domains, new drug markers, new origins of replication, or specialized functions for mapping or sequencing DNA. It should be noted that the cointegrate molecule contains Segment D (Destination vector) adjacent to segment A (Insert), thereby juxtaposing functional elements in D with the insert in A. Such molecules can be used directly in vitro (e.g., if a promoter is positioned adjacent to a gene-for in vitro transcription/translation) or in vivo (following isolation in a cell capable of propagating ccdB-containing vectors) by selecting for the selection markers in Segments B+D. As one skilled in the art will recognize, this single step method has utility in certain envisioned applications of the invention.

Figure 2 is a more detailed depiction of the recombinational cloning system of the invention, referred to herein as the "GATEWAY™ Cloning System." This figure depicts the production of Expression Clones via a "Destination Reaction," which may also be referred to herein as an "LR Reaction." A kan<sup>r</sup> vector (referred to herein as an "Entry clone") containing a DNA molecule of interest (e.g., a gene) localized between an *att*L1 site and an *att*L2 site is reacted with an amp<sup>r</sup> vector (referred to herein as a "Destination Vector") containing a toxic or "death" gene localized between an *att*R1 site and an *att*R2 site, in the presence of GATEWAY™ LR Clonase™ Enzyme Mix (a mixture of Int, IHF and Xis). After incubation at 25 °C for about 60 minutes, the reaction yields an amp<sup>r</sup> Expression Clone containing the DNA molecule of interest localized between an *att*B1 site and an *att*B2 site, and a kan<sup>r</sup> byproduct molecule, as well as intermediates. The reaction mixture may then be transformed into host cells (e.g., *E. coli*) and clones containing the nucleic acid molecule of interest may

be selected by plating the cells onto ampicillin-containing media and picking amp<sup>r</sup> colonies.

5       **Figure 3** is a schematic depiction of the cloning of a nucleic acid molecule from an Entry clone into multiple types of Destination vectors, to produce a variety of Expression Clones. Recombination between a given Entry clone and different types of Destination vectors (not shown), via the LR Reaction depicted in Figure 2, produces multiple different Expression Clones for use in a variety of applications and host cell types.

10      **Figure 4** is a detailed depiction of the production of Entry Clones via a "BP reaction," also referred to herein as an "Entry Reaction" or a "Gateward Reaction." In the example shown in this figure, an amp<sup>r</sup> expression vector containing a DNA molecule of interest (e.g., a gene) localized between an *attB1* site and an *attB2* site is reacted with a kan<sup>r</sup> Donor vector (e.g., an attP vector; here, GATEWAY™ pDONR201 (see Figure 49A-C)) containing a toxic or "death" gene localized between an *attP1* site and an *attP2* site, in the presence of GATEWAY™ BP Clonase™ Enzyme Mix (a mixture of Int and IHF). After incubation at 25°C for about 60 minutes, the reaction yields a kan<sup>r</sup> Entry clone containing the DNA molecule of interest localized between an *attL1* site and an *attL2* site, and an amp<sup>r</sup> by-product molecule. The Entry clone may then be transformed into host cells (e.g., *E. coli*) and clones containing the Entry clone (and therefore the nucleic acid molecule of interest) may be selected by plating the cells onto kanamycin-containing media and picking kan<sup>r</sup> colonies. Although this figure shows an example of use of a kan<sup>r</sup> Donor vector, it is also possible to use Donor vectors containing other selection markers, such as the gentamycin resistance or tetracycline resistance markers, as discussed herein.

15

20

25

**Figure 5** is a more detailed schematic depiction of the LR ("Destination") reaction (Figure 5A) and the BP ("Entry" or "Gateward") reaction (Figure 5B) of the GATEWAY™ Cloning System, showing the reactants, products and byproducts of each reaction.

Figure 6 shows the sequences of the attB1 and attB2 sites flanking a gene of interest after subcloning into a Destination Vector to create an Expression Clone.

Figure 7 is a schematic depiction of four ways to make Entry Clones using the compositions and methods of the invention: 1. using restriction enzymes and ligase; 2. starting with a cDNA library prepared in an attL Entry Vector; 3. using an Expression Clone from a library prepared in an attB Expression Vector via the BxP reaction; and 4. recombinational cloning of PCR fragments with terminal attB sites, via the BxP reaction. Approaches 3 and 4 rely on recombination with a Donor vector (here, an attP vector such as pDONR201 (see Figure 49A-C), pDONR202 (see Figure 50A-C), pDONR203 (see Figure 51A-C), pDONR204 (see Figure 52A-C), pDONR205 (see Figure 53A-C), or pDONR206 (see Figure 54A-C), for example) that provides an Entry Clone carrying a selection marker such as kan<sup>r</sup>, gen<sup>r</sup>, tet<sup>r</sup>, or the like.

Figure 8 is a schematic depiction of cloning of a PCR product by a BxP (Entry or Gateward) reaction. A PCR product with 25 bp terminal attB sites (plus four Gs) is shown as a substrate for the BxP reaction. Recombination between the attB-PCR product of a gene and a Donor vector (which donates an Entry Vector that carries kan<sup>r</sup>) results in an Entry Clone of the PCR product.

Figure 9 is a listing of the nucleotide sequences of the recombination sites designated herein as attB1, attB2, attP1, attP2, attL1, attL2, attR1 and attR2. Sequences are written conventionally, from 5' to 3'.

Figures 10-20: The plasmid backbone for all the Entry Vectors depicted herein is the same, and is shown in Figure 10A for the Entry Vector pENTR1A. For other Entry Vectors shown in Figures 11-20, only the sequences shown in Figure "A" for each figure set (*i.e.*, Figure 11A, Figure 12A, etc.) are different (within the attL1-attL2 cassettes) from those shown in Figure 10 -- the plasmid backbone is identical.

Figure 10 is a schematic depiction of the physical map and cloning sites (Figure 10A), and the nucleotide sequence (Figure 10B), of the Entry Vector pENTR1A.

5           **Figure 11** is a schematic depiction of the cloning sites (Figure 11A) and the nucleotide sequence (Figure 11B) of the Entry Vector pENTR2B.

10           **Figure 12** is a schematic depiction of the cloning sites (Figure 12A) and the nucleotide sequence (Figure 12B) of the Entry Vector pENTR3C.

15           **Figure 13** is a schematic depiction of the cloning sites (Figure 13A) and the nucleotide sequence (Figure 13B) of the Entry Vector pENTR4.

20           **Figure 14** is a schematic depiction of the cloning sites (Figure 14A) and the nucleotide sequence (Figure 14B) of the Entry Vector pENTR5.

25           **Figure 15** is a schematic depiction of the cloning sites (Figure 15A) and the nucleotide sequence (Figure 15B) of the Entry Vector pENTR6.

30           **Figure 16** is a schematic depiction of the cloning sites (Figure 16A) and the nucleotide sequence (Figure 16B) of the Entry Vector pENTR7.

35           **Figure 17** is a schematic depiction of the cloning sites (Figure 17A) and the nucleotide sequence (Figure 17B) of the Entry Vector pENTR8.

40           **Figure 18** is a schematic depiction of the cloning sites (Figure 18A) and the nucleotide sequence (Figure 18B) of the Entry Vector pENTR9.

45           **Figure 19** is a schematic depiction of the cloning sites (Figure 19A) and the nucleotide sequence (Figure 19B) of the Entry Vector pENTR10.

50           **Figure 20** is a schematic depiction of the cloning sites (Figure 20A) and the nucleotide sequence (Figure 20B) of the Entry Vector pENTR11.

55           **Figure 21** is a schematic depiction of the physical map and the Trc expression cassette (Figure 21A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 21B-D), of Destination Vector pDEST1. This vector may also be referred to as pTrc-DEST1.

60           **Figure 22** is a schematic depiction of the physical map and the His6 expression cassette (Figure 22A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 22B-D), of Destination Vector pDEST2. This vector may also be referred to as pHis6-DEST2.

-21-

5           **Figure 23** is a schematic depiction of the physical map and the GST expression cassette (Figure 23A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 23B-D), of Destination Vector pDEST3. This vector may also be referred to as pGST-DEST3.

10.          **Figure 24** is a schematic depiction of the physical map and the His6-Trx expression cassette (Figure 24A) showing the promoter sequences at -35 and at -10 from the initiation codon and a TEV protease cleavage site, and the nucleotide sequence (Figure 24B-D), of Destination Vector pDEST4. This vector may also be referred to as pTrx-DEST4.

15          **Figure 25** is a schematic depiction of the attR1 and attR2 sites (Figure 25A), the physical map (Figure 25B), and the nucleotide sequence (Figure 25C-D), of Destination Vector pDEST5. This vector may also be referred to as pSPORT(+-)-DEST5.

20          **Figure 26** is a schematic depiction of the attR1 and attR2 sites (Figure 26A), the physical map (Figure 26B), and the nucleotide sequence (Figure 26C-D), of Destination Vector pDEST6. This vector may also be referred to as pSPORT(-)-DEST6.

25          **Figure 27** is a schematic depiction of the attR1 site, CMV promoter, and the physical map (Figure 27A), and the nucleotide sequence (Figure 27B-C), of Destination Vector pDEST7. This vector may also be referred to as pCMV-DEST7.

30          **Figure 28** is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, and the physical map (Figure 28A), and the nucleotide sequence (Figure 28B-D), of Destination Vector pDEST8. This vector may also be referred to as pFastBac-DEST8.

30          **Figure 29** is a schematic depiction of the attR1 site, Semliki Forest Virus promoter, and the physical map (Figure 29A), and the nucleotide sequence (Figure 29B-E), of Destination Vector pDEST9. This vector may also be referred to as pSFV-DEST9.

-22-

5           **Figure 30** is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, His6 fusion domain, and the physical map (Figure 30A), and the nucleotide sequence (Figure 30B-D), of Destination Vector pDEST10. This vector may also be referred to as pFastBacHT-DEST10.

10           **Figure 31** is a schematic depiction of the attR1 cassette containing a tetracycline-regulated CMV promoter and the physical map (Figure 31A), and the nucleotide sequence (Figure 31B-D), of Destination Vector pDEST11. This vector may also be referred to as pTet-DEST11.

15           **Figure 32** is a schematic depiction of the attR1 site, the start of the mRNA of the CMV promoter, and the physical map (Figure 32A), and the nucleotide sequence (Figure 32B-D), of Destination Vector pDEST12.2. This vector may also be referred to as pCMVneo-DEST12, as pCMV-DEST12, or as pDEST12.

20           **Figure 33** is a schematic depiction of the attR1 site, the  $\lambda P_L$  promoter, and the physical map (Figure 33A), and the nucleotide sequence (Figure 33B-C), of Destination Vector pDEST13. This vector may also be referred to as p $\lambda P_L$ -DEST13.

25           **Figure 34** is a schematic depiction of the attR1 site, the T7 promoter, and the physical map (Figure 34A), and the nucleotide sequence (Figure 34B-D), of Destination Vector pDEST14. This vector may also be referred to as pPT7-DEST14.

30           **Figure 35** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 35A), and the nucleotide sequence (Figure 35B-D), of Destination Vector pDEST15. This vector may also be referred to as pT7 GST-DEST15.

25           **Figure 36** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal thioredoxin fusion sequence, and the physical map (Figure 36A), and the nucleotide sequence (Figure 36B-D), of Destination Vector pDEST16. This vector may also be referred to as pT7 Trx-DEST16.

30           **Figure 37** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal His6 fusion sequence, and the physical map (Figure 37A), and the

nucleotide sequence (Figure 37B-D), of Destination Vector pDEST17. This vector may also be referred to as pT7 His-DEST17.

5           **Figure 38** is a schematic depiction of the attR1 site and the p10 baculovirus promoter, and the physical map (Figure 38A), and the nucleotide sequence (Figure 38B-D), of Destination Vector pDEST18. This vector may also be referred to as pFBp10-DEST18.

10           **Figure 39** is a schematic depiction of the attR1 site, and the 39k baculovirus promoter, and the physical map (Figure 39A), and the nucleotide sequence (Figure 39B-D), of Destination Vector pDEST19. This vector may also be referred to as pFB39k-DEST19.

15           **Figure 40** is a schematic depiction of the attR1 site, the *polh* baculovirus promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 40A), and the nucleotide sequence (Figure 40B-D), of Destination Vector pDEST20. This vector may also be referred to as pFB GST-DEST20.

20           **Figure 41** is a schematic depiction of a 2-hybrid vector with a DNA-binding domain, the attR1 site, and the ADH promoter, and the physical map (Figure 41A), and the nucleotide sequence (Figure 41B-E), of Destination Vector pDEST21. This vector may also be referred to as pDB Leu-DEST21.

25           **Figure 42** is a schematic depiction of a 2-hybrid vector with an activation domain, the attR1 site, and the ADH promoter, and the physical map (Figure 42A), and the nucleotide sequence (Figure 42B-D), of Destination Vector pDEST22. This vector may also be referred to as pPC86-DEST22.

30           **Figure 43** is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal His6 fusion sequence, and the physical map (Figure 43A), and the nucleotide sequence (Figure 43B-D), of Destination Vector pDEST23. This vector may also be referred to as pC-term-His6-DEST23.

**Figure 44** is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal GST fusion sequence, and the physical map (Figure 44A), and the nucleotide sequence (Figure 44B-D), of Destination Vector pDEST24. This vector may also be referred to as pC-term-GST-DEST24.

5           **Figure 45** is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal thioredoxin fusion sequence, and the physical map (Figure 45A), and the nucleotide sequence (Figure 45B-D), of Destination Vector pDEST25. This vector may also be referred to as pC-term-Trx-DEST25.

10           **Figure 46** is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal His6 fusion sequence, and the physical map (Figure 46A), and the nucleotide sequence (Figure 46B-D), of Destination Vector pDEST26. This vector may also be referred to as pCMV-SPneo-His-DEST26.

15           **Figure 47** is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal GST fusion sequence, and the physical map (Figure 47A), and the nucleotide sequence (Figure 47B-D), of Destination Vector pDEST27. This vector may also be referred to as pCMV-Spneo-GST-DEST27.

20           **Figure 48** is a depiction of the physical map (Figure 48A), the cloning sites (Figure 48B), and the nucleotide sequence (Figure 48C-D), for the attB cloning vector plasmid pEXP501. This vector may also be referred to equivalently herein as pCMV•SPORT6, pCMVSPORT6, and pCMVSport6.

25           **Figure 49** is a depiction of the physical map (Figure 49A), and the nucleotide sequence (Figure 49B-C), for the Donor plasmid pDONR201 which donates a kanamycin-resistant vector in the BP Reaction. This vector may also be referred to as pAttPknnr Donor Plasmid, or as pAttPkan Donor Plasmid

**Figure 50** is a depiction of the physical map (Figure 50A), and the nucleotide sequence (Figure 50B-C), for the Donor plasmid pDONR202 which donates a kanamycin-resistant vector in the BP Reaction.

25           **Figure 51** is a depiction of the physical map (Figure 51A), and the nucleotide sequence (Figure 51B-C), for the Donor plasmid pDONR203 which donates a kanamycin-resistant vector in the BP Reaction.

**Figure 52** is a depiction of the physical map (Figure 52A), and the nucleotide sequence (Figure 52B-C), for the Donor plasmid pDONR204 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 53 is a depiction of the physical map (Figure 53A), and the nucleotide sequence (Figure 53B-C), for the Donor plasmid pDONR205 which donates a tetracycline-resistant vector in the BP Reaction.

5 Figure 54 is a depiction of the physical map (Figure 54A), and the nucleotide sequence (Figure 54B-C), for the Donor plasmid pDONR206 which donates a gentamycin-resistant vector in the BP Reaction. This vector may also be referred to as pENTR22 attP Donor Plasmid, pAttPGenr Donor Plasmid, or pAttPgnt Donor Plasmid.

10 Figure 55 depicts the attB1 site, and the physical map, of an Entry Clone (pENTR7) of CAT subcloned into the Destination Vector pDEST2 (Figure 22).

Figure 56 depicts the DNA components of Reaction B of the one-tube BxP reaction described in Example 16, pEYC7102 and attB-tet-PCR.

15 Figure 57 is a physical map of the desired product of Reaction B of the one-tube BxP reaction described in Example 16, tetx7102.

Figure 58 is a physical map of the Destination Vector pEYC8402.

20 Figure 59 is a physical map of the expected tet<sup>r</sup> subclone product, tetx8402, resulting from the LxR Reaction with tetx7102 (Figure 57) plus pEYC8402 (Figure 58).

Figure 60 is a schematic depiction of the bacteriophage lambda recombination pathways in *E. coli*.

25 Figure 61 is a schematic depiction of the DNA molecules participating in the LR Reaction. Two different co-integrates form during the LR Reaction (only one of which is shown here), depending on whether attL1 and attR1 or attL2 and attR2 are first to recombine. In one aspect, the invention provides directional cloning of a nucleic acid molecule of interest, since the recombination sites react with specificity (attL1 reacts with attR1; attL2 with attR2; attB1 with attP1; and attB2 with attP2). Thus, positioning of the sites allows construction of desired vectors having recombinant fragments in the desired orientation.

30 Figure 62 is a depiction of native and fusion protein expression using the recombinational cloning methods and compositions of the invention. In the upper figure depicting native protein expression, all of the translational start signals are

-26-

included between the attB1 and attB2 sites; therefore, these signals must be present in the starting Entry Clone. The lower figure depicts fusion protein expression (here showing expression with both N-terminal and C-terminal fusion tags so that ribosomes read through attB1 and attB2 to create the fusion protein).  
5 Unlike native protein expression vectors, N-terminal fusion vectors have their translational start signals upstream of the attB1 site.

10 **Figure 63** is a schematic depiction of three GATEWAY™ Cloning System cassettes. Three blunt-ended cassettes are depicted which convert standard expression vectors to Destination Vectors. Each of the depicted cassettes provides amino-terminal fusions in one of three possible reading frames, and each has a distinctive restriction cleavage site as shown.

15 **Figure 64** shows the physical maps of plasmids containing three attR reading frame cassettes, pEYC15101 (reading frame A; Figure 64A), pEYC15102 (reading frame B; Figure 64B), and pEYC15103 (reading frame C; Figure 64C).

20 **Figure 65** depicts the attB primers used for amplifying the tet<sup>r</sup> and amp<sup>r</sup> genes from pBR322 by the cloning methods of the invention.

25 **Figure 66** is a table listing the results of recombinational cloning of the tet<sup>r</sup> and amp<sup>r</sup> PCR products made using the primers shown in Figure 65.

30 **Figure 67** is a graph showing the effect of the number of guanines (G's) contained on the 5' end of the PCR primers on the cloning efficiency of PCR products. It is noted, however, that other nucleotides besides guanine (including A, T, C, U or combinations thereof) may be used as 5' extensions on the PCR primers to enhance cloning efficiency of PCR products.

25 **Figure 68** is a graph showing a titration of various amounts of attP and attB reactants in the BxP reaction, and the effects on cloning efficiency of PCR products.

30 **Figure 69** is a series of graphs showing the effects of various weights (Figure 69A) or moles (Figure 69B) of a 256 bp PCR product on formation of colonies, and on efficiency of cloning of the 256 bp PCR product into a Donor Vector (Figure 69C).

-27-

5           **Figure 70** is a series of graphs showing the effects of various weights (Figure 70A) or moles (Figure 70B) of a 1 kb PCR product on formation of colonies, and on efficiency of cloning of the 1 kb PCR product into a Donor Vector (Figure 70C).

10           **Figure 71** is a series of graphs showing the effects of various weights (Figure 71A) or moles (Figure 71B) of a 1.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 1.4 kb PCR product into a Donor Vector (Figure 71C).

15           **Figure 72** is a series of graphs showing the effects of various weights (Figure 72A) or moles (Figure 72B) of a 3.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 3.4 kb PCR product into a Donor Vector (Figure 72C).

20           **Figure 73** is a series of graphs showing the effects of various weights (Figure 73A) or moles (Figure 73B) of a 4.6 kb PCR product on formation of colonies, and on efficiency of cloning of the 4.6 kb PCR product into a Donor Vector (Figure 73C).

25           **Figure 74** is photograph of an ethidium bromide-stained gel of a titration of a 6.9 kb PCR product in a BxP reaction.

30           **Figure 75** is a graph showing the effects of various amounts of a 10.1 kb PCR product on formation of colonies upon cloning of the 10.1 kb PCR product into a Donor Vector.

**Figure 76** is photograph of an ethidium bromide-stained gel of a titration of a 10.1 kb PCR product in a BxP reaction.

**Figure 77** is a table summarizing the results of the PCR product cloning efficiency experiments depicted in Figures 69-74, for PCR fragments ranging in size from 0.256 kb to 6.9 kb.

**Figure 78** is a depiction of the sequences at the ends of attR Cassettes. Sequences contributed by the Cm<sup>r</sup>-ccdB cassette are shown, including the outer ends of the flanking attR sites (boxed). The staggered cleavage sites for Int are indicated in the boxed regions. Following recombination with an Entry Clone, only the outer sequences in attR sites contribute to the resulting attB sites in the

Expression Clone. The underlined sequences at both ends dictate the different reading frames (reading frames A, B, or C, with two alternative reading frame C cassettes depicted) for fusion proteins.

5           **Figure 79** is a depiction of several different attR cassettes (in reading frames A, B, or C) which may provide fusion codons at the amino-terminus of the encoded protein.

10           **Figure 80** illustrates the single-cutting restriction sites in an attR reading frame A cassette of the invention.

15           **Figure 81** illustrates the single-cutting restriction sites in an attR reading frame B cassette of the invention.

20           **Figure 82** illustrates the single-cutting restriction sites in two alternative attR reading frame C cassettes of the invention (Figures 82A and 82B) depicted in Figure 78.

25           **Figure 83** shows the physical map (Figure 83A), and the nucleotide sequence (Figure 83B-C), for an attR reading frame C parent plasmid prfC Parent III, which contains an attR reading frame C cassette of the invention (alternative A in Figures 78 and 82).

30           **Figure 84** is a physical map of plasmid pEZC1301.

20           **Figure 85** is a physical map of plasmid pEZC1313.

25           **Figure 86** is a physical map of plasmid pEZ14032.

30           **Figure 87** is a physical map of plasmid pMAB58.

35           **Figure 88** is a physical map of plasmid pMAB62.

40           **Figure 89** is a depiction of a synthesis reaction using two pairs of homologous primers of the invention.

45           **Figure 90** is a schematic depiction of the physical map (Figure 90A), and the nucleotide sequence (Figure 90B-D), of Destination Vector pDEST28.

50           **Figure 91** is a schematic depiction of the physical map (Figure 91A), and the nucleotide sequence (Figure 91B-D), of Destination Vector pDEST29.

55           **Figure 92** is a schematic depiction of the physical map (Figure 92A), and the nucleotide sequence (Figure 92B-D), of Destination Vector pDEST30.

5           **Figure 93** is a schematic depiction of the physical map (Figure 93A), and the nucleotide sequence (Figure 93B-D), of Destination Vector pDEST31.

10           **Figure 94** is a schematic depiction of the physical map (Figure 94A), and the nucleotide sequence (Figure 94B-E), of Destination Vector pDEST32.

15           **Figure 95** is a schematic depiction of the physical map (Figure 95A), and the nucleotide sequence (Figure 95B-D), of Destination Vector pDEST33.

20           **Figure 96** is a schematic depiction of the physical map (Figure 96A), and the nucleotide sequence (Figure 96B-D), of Destination Vector pDEST34.

25           **Figure 97** is a depiction of the physical map (Figure 97A), and the nucleotide sequence (Figure 97B-C), for the Donor plasmid pDONR207 which donates a gentamycin-resistant vector in the BP Reaction.

30           **Figure 98** is a schematic depiction of the physical map (Figure 98A), and the nucleotide sequence (Figure 98B-D), of the 2-hybrid vector pMAB85.

35           **Figure 99** is a schematic depiction of the physical map (Figure 99A), and the nucleotide sequence (Figure 99B-D), of the 2-hybrid vector pMAB86.

## DETAILED DESCRIPTION OF THE INVENTION

### 20           *Definitions*

In the description that follows, a number of terms used in recombinant DNA technology are utilized extensively. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

25           **Byproduct:** is a daughter molecule (a new clone produced after the second recombination event during the recombinational cloning process) lacking the segment which is desired to be cloned or subcloned.

30           **Cointegrate:** is at least one recombination intermediate nucleic acid molecule of the present invention that contains both parental (starting) molecules. It will usually be linear. In some embodiments it can be circular. RNA and polypeptides may be expressed from cointegrates using an appropriate host cell strain, for example *E. coli* DB3.1 (particularly *E. coli* LIBRARY EFFICIENCY®

DB3.1<sup>TM</sup> Competent Cells), and selecting for both selection markers found on the cointegrate molecule.

5           **Host:** is any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Product, vector, or nucleic acid molecule of the invention. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of such hosts, see Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982).

10           **Insert or Inserts:** include the desired nucleic acid segment or a population 10 of nucleic acid segments (segment *A* of Figure 1) which may be manipulated by the methods of the present invention. Thus, the terms Insert(s) are meant to include a particular nucleic acid (preferably DNA) segment or a population of segments. Such Insert(s) can comprise one or more nucleic acid molecules.

15           **Insert Donor:** is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the Insert. The Insert Donor molecule comprises the Insert flanked on both sides with recombination sites. The Insert Donor can be linear or circular. In one embodiment of the invention, 20 the Insert Donor is a circular DNA molecule and further comprises a cloning vector sequence outside of the recombination signals (see Figure 1). When a population of Inserts or population of nucleic acid segments are used to make the Insert Donor, a population of Insert Donors results and may be used in accordance with the invention. Examples of such Insert Donor molecules are GATEWAY<sup>TM</sup> Entry Vectors, which include but are not limited to those Entry Vectors depicted in Figures 10-20, as well as other vectors comprising a gene of interest flanked by 25 one or more *attL* sites (e.g., *attL1*, *attL2*, etc.), or by one or more *attB* sites (e.g., *attB1*, *attB2*, etc.) for the production of library clones.

30           **Product:** is one of the desired daughter molecules comprising the *A* and *D* sequences which is produced after the second recombination event during the recombinational cloning process (see Figure 1). The Product contains the nucleic acid which was to be cloned or subcloned. In accordance with the invention, when a population of Insert Donors are used, the resulting population of Product

molecules will contain all or a portion of the population of Inserts of the Insert Donors and preferably will contain a representative population of the original molecules of the Insert Donors.

**Promoter:** is a DNA sequence generally described as the 5'-region of a gene, located proximal to the start codon. The transcription of an adjacent DNA segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

**Recognition sequence:** Recognition sequences are particular sequences which a protein, chemical compound, DNA, or RNA molecule (*e.g.*, restriction endonuclease, a modification methylase, or a recombinase) recognizes and binds. In the present invention, a recognition sequence will usually refer to a recombination site. For example, the recognition sequence for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. *See Figure 1 of Sauer, B., Current Opinion in Biotechnology 5:521-527 (1994).* Other examples of recognition sequences are the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombinase enzyme  $\lambda$  Integrase. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). *See Landy, Current Opinion in Biotechnology 3:699-707 (1993).* Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (*e.g.*, *attR'* or *attP'*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way.

5           **Recombination proteins:** include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites, which may be wild-type proteins (*See Landy, Current Opinion in Biotechnology 3:699-707 (1993)*), or mutants, derivatives (e.g., fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof.

10           **Recombination site:** is a recognition sequence on a DNA molecule participating in an integration/recombination reaction by the recombinational cloning methods of the invention. Recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. *See Figure 1 of Sauer, B., Curr. Opin. Biotech. 5:521-527 (1994).* Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences described herein, and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein  $\lambda$  Int and by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). *See Landy, Curr. Opin. Biotech. 3:699-707 (1993).*

15

20

25           **Recombinational Cloning:** is a method described herein, whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. By “*in vitro*” and “*in vivo*” herein is meant recombinational cloning that is carried out outside of host cells (e.g., in cell-free systems) or inside of host cells (e.g., using recombination proteins expressed by host cells), respectively.

30           **Repression cassette:** is a nucleic acid segment that contains a repressor or a Selectable marker present in the subcloning vector.

35           **Selectable marker:** is a DNA segment that allows one to select for or against a molecule (e.g., a replicon) or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to,

production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of Selectable markers include but are not limited to: (1) DNA segments that encode products which provide resistance against otherwise toxic compounds (e.g., antibiotics); (2) DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); (3) DNA segments that encode products which suppress the activity of a gene product; (4) DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as  $\beta$ -galactosidase, green fluorescent protein (GFP), and cell surface proteins); (5) DNA segments that bind products which are otherwise detrimental to cell survival and/or function; (6) DNA segments that otherwise inhibit the activity of any of the DNA segments described in Nos. 1-5 above (e.g., antisense oligonucleotides); (7) DNA segments that bind products that modify a substrate (e.g. restriction endonucleases); (8) DNA segments that can be used to isolate or identify a desired molecule (e.g. specific protein binding sites); (9) DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (e.g., for PCR amplification of subpopulations of molecules); (10) DNA segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; (11) DNA segments that encode products which are toxic in recipient cells; (12) DNA segments that inhibit replication, partition or heritability of nucleic acid molecules that contain them; and/or (13) DNA segments that encode conditional replication functions, e.g., replication in certain hosts or host cell strains or under certain environmental conditions (e.g., temperature, nutritional conditions, etc.).

**Selection scheme:** is any method which allows selection, enrichment, or identification of a desired Product or Product(s) from a mixture containing an Entry Clone or Vector, a Destination Vector, a Donor Vector, an Expression Clone or Vector, any intermediates (e.g. a Cointegrate or a replicon), and/or Byproducts. The selection schemes of one preferred embodiment have at least two components that are either linked or unlinked during recombinational cloning. One component is a Selectable marker. The other component controls the expression *in vitro* or *in vivo* of the Selectable marker, or survival of the cell (or

the nucleic acid molecule, e.g., a replicon) harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression or activity of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various DNA segments, as will be readily apparent to those skilled in the art. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or more desired Products. As defined herein, selecting for a DNA molecule includes (a) selecting or enriching for the presence of the desired DNA molecule, and (b) selecting or enriching against the presence of DNA molecules that are not the desired DNA molecule.

In one embodiment, the selection schemes (which can be carried out in reverse) will take one of three forms, which will be discussed in terms of Figure 1. The first, exemplified herein with a Selectable marker and a repressor therefore, selects for molecules having segment *D* and lacking segment *C*. The second selects against molecules having segment *C* and for molecules having segment *D*. Possible embodiments of the second form would have a DNA segment carrying a gene toxic to cells into which the *in vitro* reaction products are to be introduced. A toxic gene can be a DNA that is expressed as a toxic gene product (a toxic protein or RNA), or can be toxic in and of itself. (In the latter case, the toxic gene is understood to carry its classical definition of "heritable trait".)

Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (e.g., *Dpn*I), apoptosis-related genes (e.g. ASK1 or members of the bcl-2/ced-9 family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic DNA sequences, bacteriophage lytic genes such as those from  $\Phi$ X174 or bacteriophage T4; antibiotic sensitivity genes such as *rpsL*, antimicrobial sensitivity genes such as *pheS*, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the absence of a suppressing function, e.g., *kicB*, *ccdB*,  $\Phi$ X174 E (Liu, Q. et al., *Curr. Biol.*

8:1300-1309 (1998)), and other genes that negatively affect replicon stability and/or replication. A toxic gene can alternatively be selectable *in vitro*, e.g., a restriction site.

Many genes coding for restriction endonucleases operably linked to inducible promoters are known, and may be used in the present invention. See, e.g. U.S. Patent Nos. 4,960,707 (*DpnI* and *DpnII*); 5,000,333, 5,082,784 and 5,192,675 (*KpnI*); 5,147,800 (*NgoAIII* and *NgoAI*); 5,179,015 (*FspI* and *HaeIII*); 5,200,333 (*HaeII* and *TaqI*); 5,248,605 (*HpaII*); 5,312,746 (*Clal*); 5,231,021 and 5,304,480 (*XhoI* and *XhoII*); 5,334,526 (*AluI*); 5,470,740 (*NsiI*); 5,534,428 (*SstI/SacI*); 5,202,248 (*NcoI*); 5,139,942 (*NdeI*); and 5,098,839 (*PacI*). See also Wilson, G.G., *Nucl. Acids Res.* 19:2539-2566 (1991); and Lunnen, K.D., *et al.*, *Gene* 74:25-32 (1988).

In the second form, segment *D* carries a Selectable marker. The toxic gene would eliminate transformants harboring the Vector Donor, Cointegrate, and Byproduct molecules, while the Selectable marker can be used to select for cells containing the Product and against cells harboring only the Insert Donor.

The third form selects for cells that have both segments *A* and *D* in *cis* on the same molecule, but not for cells that have both segments in *trans* on different molecules. This could be embodied by a Selectable marker that is split into two inactive fragments, one each on segments *A* and *D*.

The fragments are so arranged relative to the recombination sites that when the segments are brought together by the recombination event, they reconstitute a functional Selectable marker. For example, the recombinational event can link a promoter with a structural nucleic acid molecule (e.g., a gene), can link two fragments of a structural nucleic acid molecule, or can link nucleic acid molecules that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

**Site-specific recombinase:** is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase

activity to reseal the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of sequence specificity for both partners. The strand exchange mechanism involves the cleavage and rejoicing of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

5

**Subcloning vector:** is a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector (segment *D* in Figure 1) can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned DNA Insert (segment *A* in Figure 1). The subcloning vector can also contain a Selectable marker (preferably DNA).

10

**Vector:** is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Patent No. 5,334,575, entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

15

20

25

30

5           **Vector Donor:** is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the DNA segments comprising the DNA vector which is to become part of the desired Product. The Vector Donor comprises a subcloning vector *D* (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vector (e.g., for PCR fragments containing *attB* sites; see below)) and a segment *C* flanked by recombination sites (see Figure 1). Segments *C* and/or *D* can contain elements that contribute to selection for the desired Product daughter molecule, as described above for selection schemes. The recombination signals can be the same or different, and can be acted upon by the same or different recombinases. In  
10           addition, the Vector Donor can be linear or circular. Examples of such Vector Donor molecules include GATEWAY™ Destination Vectors, which include but are not limited to those Destination Vectors depicted in Figures 21-47 and 90-96.

15           **Primer:** refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (e.g. a DNA molecule). In a preferred aspect, a primer comprises one or more recombination sites or portions of such recombination sites. Portions of recombination sites comprise at least 2 bases (or basepairs, abbreviated herein as "bp"), at least 5-200 bases, at least 10-100 bases,  
20           at least 15-75 bases, at least 15-50 bases, at least 15-25 bases, or at least 16-25 bases, of the recombination sites of interest, as described in further detail below and in the Examples. When using portions of recombination sites, the missing portion of the recombination site may be provided as a template by the newly synthesized nucleic acid molecule. Such recombination sites may be located within and/or at one or both termini of the primer. Preferably, additional sequences are added to the primer adjacent to the recombination site(s) to enhance or improve recombination and/or to stabilize the recombination site during recombination. Such stabilization sequences may be any sequences (preferably G/C rich sequences) of any length. Preferably, such sequences range in size from  
25           1 to about 1000 bases, 1 to about 500 bases, and 1 to about 100 bases, 1 to about 60 bases, 1 to about 25, 1 to about 10, 2 to about 10 and preferably about 4 bases.  
30

Preferably, such sequences are greater than 1 base in length and preferably greater than 2 bases in length.

**Template:** refers to double stranded or single stranded nucleic acid molecules which are to be amplified, synthesized or sequenced. In the case of double stranded molecules, denaturation of its strands to form a first and a second strand is preferably performed before these molecules will be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to a portion of the template is hybridized under appropriate conditions and one or more polypeptides having polymerase activity (e.g. DNA polymerases and/or reverse transcriptases) may then synthesize a nucleic acid molecule complementary to all or a portion of said template. Alternatively, for double stranded templates, one or more promoters may be used in combination with one or more polymerases to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecules, according to the invention, may be equal or shorter in length than the original template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

**Adapter:** is an oligonucleotide or nucleic acid fragment or segment (preferably DNA) which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear Insert Donor molecule as well as other nucleic acid molecules described herein. When using portions of recombination sites, the missing portion may be provided by the Insert Donor molecule. Such adapters may be added at any location within a circular or linear molecule, although the adapters are preferably added at or near one or both termini of a linear molecule. Preferably, adapters are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adapters may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adapters may be added to a circular molecule by first digesting the molecule with

an appropriate restriction enzyme, adding the adapter at the cleavage site and reforming the circular molecule which contains the adapter(s) at the site of cleavage. In other aspects, adapters may be added by homologous recombination, by integration of RNA molecules, and the like. Alternatively, adapters may be ligated directly to one or more and preferably both termini of a linear molecule thereby resulting in linear molecule(s) having adapters at one or both termini. In one aspect of the invention, adapters may be added to a population of linear molecules, (e.g. a cDNA library or genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adapters at one and preferably both termini of all or substantial portion of said population.

**Adapter-Primer:** is primer molecule which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear nucleic acid molecule described herein. When using portions of recombination sites, the missing portion may be provided by a nucleic acid molecule (e.g., an adapter) of the invention. Such adapter-primers may be added at any location within a circular or linear molecule, although the adapter-primers are preferably added at or near one or both termini of a linear molecule. Examples of such adapter-primers and the use thereof in accordance with the methods of the invention are shown in Example 25 herein. Such adapter-primers may be used to add one or more recombination sites or portions thereof to circular or linear nucleic acid molecules in a variety of contexts and by a variety of techniques, including but not limited to amplification (e.g., PCR), ligation (e.g., enzymatic or chemical/synthetic ligation), recombination (e.g., homologous or non-homologous (illegitimate) recombination) and the like.

**Library:** refers to a collection of nucleic acid molecules (circular or linear). In one embodiment, a library may comprise a plurality (*i.e.*, two or more) of DNA molecules, which may or may not be from a common source organism, organ, tissue, or cell. In another embodiment, a library is representative of all or a portion or a significant portion of the DNA content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a portion or a significant portion of the expressed nucleic acid molecules (a cDNA library) in a

cell, tissue, organ or organism. A library may also comprise random sequences made by *de novo* synthesis, mutagenesis of one or more sequences and the like. Such libraries may or may not be contained in one or more vectors.

**Amplification:** refers to any *in vitro* method for increasing a number of copies of a nucleotide sequence with the use of a polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5-100 "cycles" of denaturation and synthesis of a DNA molecule.

**Oligonucleotide:** refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides which are joined by a phosphodiester bond between the 3' position of the deoxyribose or ribose of one nucleotide and the 5' position of the deoxyribose or ribose of the adjacent nucleotide. This term may be used interchangeably herein with the terms "nucleic acid molecule" and "polynucleotide," without any of these terms necessarily indicating any particular length of the nucleic acid molecule to which the term specifically refers.

**Nucleotide:** refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [ $\alpha$ S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

5                   **Hybridization:** The terms "hybridization" and "hybridizing" refers to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may be hybridized, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used. In some aspects, hybridization is said to be under "stringent conditions." By "stringent conditions" as used herein is meant overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

10

15                  Other terms used in the fields of recombinant DNA technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.

### *Overview*

20                  Two reactions constitute the recombinational cloning system of the present invention, referred to herein as the "GATEWAY™ Cloning System," as depicted generally in Figure 1. The first of these reactions, the **LR Reaction** (Figure 2), which may also be referred to interchangeably herein as the **Destination Reaction**, is the main pathway of this system. The LR Reaction is a recombination reaction between an Entry vector or clone and a Destination Vector, mediated by a cocktail of recombination proteins such as the GATEWAY™ LR Clonase™ Enzyme Mix described herein. This reaction transfers nucleic acid molecules of interest (which may be genes, cDNAs, cDNA libraries, or fragments thereof) from the Entry Clone to an Expression Vector, to create an Expression Clone.

25

30                  The sites labeled L, R, B, and P are respectively the attL, attR, attB, and attP recombination sites for the bacteriophage λ recombination proteins that constitute the Clonase cocktail (referred to herein variously as "Clonase" or

“GATEWAY™ LR Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attL x attR recombination reactions, as described herein) or “GATEWAY™ BP Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attB x attP recombination reactions, as described herein)). The Recombinational Cloning reactions are equivalent to concerted, highly specific, cutting and ligation reactions. Viewed in this way, the recombination proteins cut to the left and right of the nucleic acid molecule of interest in the Entry Clone and ligate it into the Destination vector, creating a new Expression Clone.

The nucleic acid molecule of interest in an Expression Clone is flanked by the small attB1 and attB2 sites. The orientation and reading frame of the nucleic acid molecule of interest are maintained throughout the subcloning, because attL1 reacts only with attR1, and attL2 reacts only with attR2. Likewise, attB1 reacts only with attP1, and attB2 reacts only with attP2. Thus, the invention also relates to methods of controlled or directional cloning using the recombination sites of the invention (or portions thereof), including variants, fragments, mutants and derivatives thereof which may have altered or enhanced specificity. The invention also relates more generally to any number of recombination site partners or pairs (where each recombination site is specific for and interacts with its corresponding recombination site). Such recombination sites are preferably made by mutating or modifying the recombination site to provide any number of necessary specificities (e.g., attB1-10, attP1-10, attL1-10, attR1-10, etc.), non-limiting examples of which are described in detail in the Examples herein.

When an aliquot from the recombination reaction is transformed into host cells (e.g., *E. coli*) and spread on plates containing an appropriate selection agent, e.g., an antibiotic such as ampicillin with or without methicillin, cells that take up the desired clone form colonies. The unreacted Destination Vector does not give ampicillin-resistant colonies, even though it carries the ampicillin-resistance gene, because it contains a toxic gene, e.g., *ccdB*. Thus selection for ampicillin resistance selects for *E. coli* cells that carry the desired product, which usually comprise >90% of the colonies on the ampicillin plate.

To participate in the Recombinational (or “GATEWAY™”) Cloning Reaction, a nucleic acid molecule of interest first may be cloned into an Entry

Vector, creating an Entry Clone. Multiple options are available for creating Entry Clones, including: cloning of PCR sequences with terminal attB recombination sites into Entry Vectors; using the GATEWAY™ Cloning System recombination reaction; transfer of genes from libraries prepared in GATEWAY™ Cloning System vectors by recombination into Entry Vectors; and cloning of restriction enzyme-generated fragments and PCR fragments into Entry Vectors by standard recombinant DNA methods. These approaches are discussed in further detail herein.

A key advantage of the GATEWAY™ Cloning System is that a nucleic acid molecule of interest (or even a population of nucleic acid molecules of interest) present as an Entry Clone can be subcloned in parallel into one or more Destination Vectors in a simple reactions for anywhere from about 30 seconds to about 60 minutes (preferably about 1-60 minutes, about 1-45 minutes, about 1-30 minutes, about 2-60 minutes, about 2-45 minutes, about 2-30 minutes, about 1-2 minutes, about 30-60 minutes, about 45-60 minutes, or about 30-45 minutes). Longer reaction times (*e.g.*, 2-24 hours, or overnight) may increase recombination efficiency, particularly where larger nucleic acid molecules are used, as described in the Examples herein. Moreover, a high percentage of the colonies obtained carry the desired Expression Clone. This process is illustrated schematically in Figure 3, which shows an advantage of the invention in which the molecule of interest can be moved simultaneously or separately into multiple Destination Vectors. In the LR Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (*e.g.*, linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

The second major pathway of the GATEWAY™ Cloning System is the **BP Reaction** (Figure 4), which may also be referred to interchangeably herein as the **Entry Reaction** or the **Gateway Reaction**. The BP Reaction may recombine an Expression Clone with a Donor Plasmid (the counterpart of the byproduct in Figure 2). This reaction transfers the nucleic acid molecule of interest (which may have any of a variety of topologies, including linear, coiled, supercoiled, etc.) in the Expression Clone into an Entry Vector, to produce a new Entry Clone. Once this nucleic acid molecule of interest is cloned into an Entry

Vector, it can be transferred into new Expression Vectors, through the LR Reaction as described above. In the BP Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (*e.g.*, linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

5 A useful variation of the BP Reaction permits rapid cloning and expression of products of amplification (*e.g.*, PCR) or nucleic acid synthesis. Amplification (*e.g.*, PCR) products synthesized with primers containing terminal 25 bp attB sites serve as efficient substrates for the Gateward Cloning reaction. Such amplification products may be recombined with a Donor Vector to produce an Entry Clone (see  
10 Figure 7). The result is an Entry Clone containing the amplification fragment. Such Entry Clones can then be recombined with Destination Vectors -- through the LR Reaction -- to yield Expression Clones of the PCR product.

15 Additional details of the LR Reaction are shown in Figure 5A. The GATEWAY™ LR Clonase™ Enzyme Mix that mediates this reaction contains lambda recombination proteins Int (Integrase), Xis (Excisionase), and IHF (Integration Host Factor). In contrast, the GATEWAY™ BP Clonase™ Enzyme Mix, which mediates the BP Reaction (Figure 5B), comprises Int and IHF alone.

20 The recombination (att) sites of each vector comprise two distinct segments, donated by the parental vectors. The staggered lines dividing the two portions of each att site, depicted in Figures 5A and 5B, represent the seven-base staggered cut produced by Int during the recombination reactions. This structure is seen in greater detail in Figure 6, which displays the attB recombination sequences of an Expression Clone, generated by recombination between the attL1 and attL2 sites of an Entry Clone and the attR1 and attR2 sites of a Destination  
25 Vector.

30 The nucleic acid molecule of interest in the Expression Clone is flanked by attB sites: attB1 to the left (amino terminus) and attB2 to the right (carboxy terminus). The bases in attB1 to the left of the seven-base staggered cut produced by Int are derived from the Destination vector, and the bases to the right of the staggered cut are derived from the Entry Vector (see Figure 6). Note that the sequence is displayed in triplets corresponding to an open reading frame. If the reading frame of the nucleic acid molecule of interest cloned in the Entry Vector

is in phase with the reading frame shown for attB1, amino-terminal protein fusions can be made between the nucleic acid molecule of interest and any GATEWAY™ Cloning System Destination Vector encoding an amino-terminal fusion domain. Entry Vectors and Destination Vectors that enable cloning in all three reading frames are described in more detail herein, particularly in the Examples.

The LR Reaction allows the transfer of a desired nucleic acid molecule of interest into new Expression Vectors by recombining a Entry Clone with various Destination Vectors. To participate in the LR or Destination Reaction, however, a nucleic acid molecule of interest preferably is first converted to a Entry Clone. Entry Clones can be made in a number of ways, as shown in Figure 7.

One approach is to clone the nucleic acid molecule of interest into one or more of the Entry Vectors, using standard recombinant DNA methods, with restriction enzymes and ligase. The starting DNA fragment can be generated by restriction enzyme digestion or as a PCR product. The fragment is cloned between the attL1 and attL2 recombination sites in the Entry Vector. Note that a toxic or "death" gene (*e.g.*, ccdB), provided to minimize background colonies from incompletely digested Entry Vector, must be excised and replaced by the nucleic acid molecule of interest.

A second approach to making an Entry Clone (Figure 7) is to make a library (genomic or cDNA) in an Entry Vector, as described in detail herein. Such libraries may then be transferred into Destination Vectors for expression screening, for example in appropriate host cells such as yeast cells or mammalian cells.

A third approach to making Entry Clones (Figure 7) is to use Expression Clones obtained from cDNA molecules or libraries prepared in Expression Vectors. Such cDNAs or libraries, flanked by attB sites, can be introduced into a Entry Vector by recombination with a Donor Vector via the BP Reaction. If desired, an entire Expression Clone library can be transferred into the Entry Vector through the BP Reaction. Expression Clone cDNA libraries may also be constructed in a variety of prokaryotic and eukaryotic GATEWAY™-modified vectors (*e.g.*, the pEXP501 Expression Vector (see Figure 48), and 2-hybrid and

attB library vectors), as described in detail herein, particularly in the Examples below.

A fourth, and potentially most versatile, approach to making an Entry Clone (Figure 7) is to introduce a sequence for a nucleic acid molecule of interest into an Entry Vector by amplification (e.g., PCR) fragment cloning. This method is diagramed in Figure 8. The DNA sequence first is amplified (for example, with PCR) as outlined in detail below and in the Examples herein, using primers containing one or more bp, two or more bp, three or more bp, four or more bp, five or more bp, preferably six or more bp, more preferably 6-25 bp (particularly 5 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25) bp of the attB nucleotide sequences (such as, but not limited to, those depicted in Figure 9), and optionally one or more, two or more, three or more, four or more, and most preferably four 10 or five or more additional terminal nucleotide bases which preferably are guanines. The PCR product then may be converted to a Entry Clone by performing a BP Reaction, in which the attB-PCR product recombines with a Donor Vector containing one or more attP sites. Details of this approach and protocols for PCR 15 fragment subcloning are provided in Examples 8 and 21-25.

A variety of Entry Clones may be produced by these methods, providing 20 a wide array of cloning options; a number of specific Entry Vectors are also available commercially from Life Technologies, Inc. (Rockville, MD). The Examples herein provide a more in-depth description of selected Entry Vectors and details of their cloning sites. Choosing the optimal Entry Vector for a particular application is discussed in Example 4.

Entry Vectors and Destination Vectors should be constructed so that the 25 amino-terminal region of a nucleic acid molecule of interest (e.g., a gene, cDNA library or insert, or fragment thereof) will be positioned next to the attL1 site. Entry Vectors preferably contain the rrnB transcriptional terminator upstream of the attL1 site. This sequence ensures that expression of cloned nucleic acid molecules of interest is reliably "off" in *E. coli*, so that even toxic genes can be successfully cloned. Thus, Entry Clones may be designed to be transcriptionally 30 silent. Note also that Entry Vectors, and hence Entry Clones, may contain the kanamycin antibiotic resistance (kan') gene to facilitate selection of host cells

containing Entry Clones after transformation. In certain applications, however, Entry Clones may contain other selection markers, including but not limited to a gentamycin resistance (*gen*<sup>r</sup>) or tetracycline resistance (*tet*<sup>r</sup>) gene, to facilitate selection of host cells containing Entry Clones after transformation.

Once a nucleic acid molecule of interest has been cloned into an Entry Vector, it may be moved into a Destination Vector. The upper right portion of Figure 5A shows a schematic of a Destination Vector. The thick arrow represents some function (often transcription or translation) that will act on the nucleic acid molecule of interest in the clone. During the recombination reaction, the region between the attR1 and attR2 sites, including a toxic or "death" gene (e.g., *ccdB*), is replaced by the DNA segment from the Entry Clone. Selection for recombinants that have acquired the ampicillin resistance (*amp*<sup>r</sup>) gene (carried on the Destination Vector) and that have also lost the death gene ensures that a high percentage (usually >90%) of the resulting colonies will contain the correct insert.

To move a nucleic acid molecule of interest into a Destination Vector, the Destination Vector is mixed with the Entry Clone comprising the desired nucleic acid molecule of interest, a cocktail of recombination proteins (e.g., GATEWAY™ LR Clonase™ Enzyme Mix) is added, the mixture is incubated (preferably at about 25°C for about 60 minutes, or longer under certain circumstances, e.g. for transfer of large nucleic acid molecules, as described below) and any standard host cell (including bacterial cells such as *E. coli*; animal cells such as insect cells, mammalian cells, nematode cells and the like; plant cells; and yeast cells) strain is transformed with the reaction mixture. The host cell used will be determined by the desired selection (e.g., *E. coli* DB3.1, available commercially from Life Technologies, Inc., allows survival of clones containing the *ccdB* death gene, and thus can be used to select for cointegrate molecules -- i.e., molecules that are hybrids between the Entry Clone and Destination Vector). The Examples below provide further details and protocols for use of Entry and Destination Vectors in transferring nucleic acid molecules of interest and expressing RNAs or polypeptides encoded by these nucleic acid molecules in a variety of host cells.

The cloning system of the invention therefore offers multiple advantages:

- Once a nucleic acid molecule of interest is cloned into the GATEWAY™ Cloning System, it can be moved into and out of other vectors with complete fidelity of reading frame and orientation. That is, since the reactions proceed whereby attL1 on the Entry Clone recombines with attR1 on the Destination Vector, the directionality of the nucleic acid molecule of interest is maintained or may be controlled upon transfer from the Entry Clone into the Destination Vector. Hence, the GATEWAY™ Cloning System provides a powerful and easy method of directional cloning of nucleic acid molecule of interest.
- One-step cloning or subcloning: Mix the Entry Clone and the Destination Vector with Clonase, incubate, and transform.
- Clone PCR products readily by *in vitro* recombination, by adding attB sites to PCR primers. Then directly transfer these Entry Clones into Destination Vectors. This process may also be carried out in one step (see Examples below).
- Powerful selections give high reliability: >90% (and often >99%) of the colonies contain the desired DNA in its new vector.
- One-step conversion of existing standard vectors into GATEWAY™ Cloning System vectors.
- Ideal for large vectors or those with few cloning sites.
- Recombination sites are short (25 bp), and may be engineered to contain no stop codons or secondary structures.
- Reactions may be automated, for high-throughput applications (*e.g.*, for diagnostic purposes or for therapeutic candidate screening).
- The reactions are economical: 0.3 µg of each DNA; no restriction enzymes, phosphatase, ligase, or gel purification. Reactions work well with miniprep DNA.
- Transfer multiple clones, and even libraries, into one or more Destination Vectors, in a single experiment.
- A variety of Destination Vectors may be produced, for applications including, but not limited to:

- Protein expression in *E. coli*: native proteins; fusion proteins with GST, His6, thioredoxin, etc., for purification, or one or more epitope tags; any promoter useful in expressing proteins in *E. coli* may be used, such as ptrc,  $\lambda P_L$ , and T7 promoters.
- 5           • Protein expression in eukaryotic cells: CMV promoter, baculovirus (with or without His6 tag), Semliki Forest virus, Tet regulation.
- DNA sequencing (all *lac* primers), RNA probes, phagemids (both strands)
- A variety of Entry Vectors (for recombinational cloning entry by standard recombinant DNA methods) may be produced:
  - Strong transcription stop just upstream, for genes toxic to *E. coli*.
  - Three reading frames.
  - With or without TEV protease cleavage site.
  - Motifs for prokaryotic and / or eukaryotic translation.
  - Compatible with commercial cDNA libraries.
- 10          • Expression Clone cDNA (attB) libraries, for expression screening, including 2-hybrid libraries and phage display libraries, may also be constructed.
- 15

### ***Recombination Site Sequences***

20          In one aspect, the invention relates to nucleic acid molecules, which may or may not be isolated nucleic acid molecules, comprising one or more nucleotide sequences encoding one or more recombination sites or portions thereof. In particular, this aspect of the invention relates to such nucleic acid molecules comprising one or more nucleotide sequences encoding *attB*, *attP*, *attL*, or *attR*, or portions of these recombination site sequences. The invention also relates to mutants, derivatives, and fragments of such nucleic acid molecules. Unless otherwise indicated, all nucleotide sequences that may have been determined by sequencing a DNA molecule herein were determined using manual or automated DNA sequencing, such as dideoxy sequencing, according to methods that are routine to one of ordinary skill in the art (Sanger, F., and Coulson, A.R., *J. Mol. Biol.* 94:444-448 (1975); Sanger, F., et al., *Proc. Natl. Acad. Sci. USA* 74:5463-5467 (1977)). All amino acid sequences of polypeptides encoded by DNA

25

30

-50-

5

10

15

20

25

30

molecules determined herein were predicted by conceptual translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by these approaches, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by such methods are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G, C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U). Thus, the invention relates to sequences of the invention in the form of DNA or RNA molecules, or hybrid DNA/RNA molecules, and their corresponding complementary DNA, RNA, or DNA/RNA strands.

In a first such aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTGTACAAAAAAAGCAGGCT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB1*, or mutants, fragments, variants or derivatives thereof. As one of ordinary skill will appreciate, however, certain mutations, insertions, or deletions of one or more bases in the *attB1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional

integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB1* sequence are encompassed within the scope of the invention.

In a related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB2* nucleotide sequence having the sequence set forth in Figure 9, such as: ACCCAGCTTCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attB2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule containing *attB1* and *attB2* sites (the vector pEXP501, also known as pCMVSport6; see Figure 48), *E. coli* DB3.1(pCMVSport6), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30108. The *attB1* and *attB2* sites within the deposited nucleic acid molecule are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP1* nucleotide sequence having the sequence set forth in Figure 9, such as: TACAGGTCACTAATACCATCTAACGTTGATTGATAGTGA-CTGGATATGTTGTGTTTACAGTATTATGATAGTCTGTTTTAT-GCAAAATCTAATTAAATATATTGATATTATATCATTTCAGTT-TCTCGTTCAAGCTTTTGACAAAGTTGGCATTATAAAAAAGCATTG-CTCATCAATTGTTGCAACGAACAGGTCACTATCAGTAAAATAA-

AATCATTATTTG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP2* nucleotide sequence having the sequence set forth in Figure 9, such as: CAAATAATGATTATTTGACTGATAGTGACCTGTTCGTTG-  
CAACAAATTGATAAGCAATGCTTCTTATAATGCCAACTTT-  
GTACAAGAAAGCTAACGAGAACGTAAGATGATA-  
TAAATATCAATATATTAAATTAGATTTCGATAAAAAACAG-  
ACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAA-  
CTACTTAGATGGTATTAGTGACCTGTA, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule (the *attP* vector pDONR201, also known as pENTR21-*attPkan* or pAttPkan; see Figure 49) containing *attP1* and *attP2* sites, *E. coli* DB3.1(pAttPkan) (also called *E. coli* DB3.1(pAHKan)), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30099. The *attP1* and *attP2* sites within the deposited nucleic acid molecule are contained in nucleic acid

cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTGTACAAAAAAGCTGAACGAG-AAACGTAAAATGATATAATATCAATATATTAAATTAGATTTGCAT-AAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCA-CTATG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR2* nucleotide sequence having the sequence set forth in Figure 9, such as: G C A G G T C G A C C A T A G T G A C T G G A T A T-GTTGTGTTTACAGTATTATGTAGTCTGTTTTATGCAAATCTA-ATTAAATATATTGATATTATCATTACGTTCTCGTTAGCTT-TCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR2* sequence are encompassed within the scope of the invention.

Recombinant host cell strains containing attR1 sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pEZR15101) (reading frame A; see Figure 64A), *E. coli* DB3.1(pEZR15102) (reading frame B; see Figure 64B), and *E. coli* DB3.1(pEZR15103) (reading frame C; see Figure 64C), and containing corresponding attR2 sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30103, NRRL B-30104, and NRRL B-30105, respectively. The attR1 and attR2 sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL1*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL1* nucleotide sequence having the sequence set forth in Figure 9, such as: CAA ATA ATG ATT TTA TTT TGA CTG ATA GTG ACC TGT TCG TTG CAA CAA ATT GAT AAG CAA TGC TTT TTT ATA ATG CCA ACT TTG TAC AAA AAA GCA GGC T, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL2*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL2* nucleotide sequence having the sequence set forth in Figure 9, such as: C AAA TAA TGA TTT TAT TTT GAC TGA TAG TGA CCT GTT CGT TGC AAC AAA TTG ATA AGC AAT GCT TTC TTA TAA TGC CAA

5 CTT TGT ACA AGA AAG CTG GGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL2* sequence are encompassed within the scope of the invention.

10 Recombinant host cell strains containing *attL1* sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pENTR1A) (reading frame A; see Figure 10), *E. coli* DB3.1(pENTR2B) (reading frame B; see Figure 11), and *E. coli* DB3.1(pENTR3C) (reading frame C; see Figure 12), and containing corresponding *attL2* sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30100, NRRL B-30101, and NRRL B-30102, respectively. The *attL1* and *attL2* sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

15

20 Each of the recombination site sequences described herein or portions thereof, or the nucleotide sequence cassettes contained in the deposited clones, may be cloned or inserted into a vector of interest (for example, using the recombinational cloning methods described herein and/or standard restriction cloning techniques that are routine in the art) to generate, for example, Entry Vectors or Destination Vectors which may be used to transfer a desired segment of a nucleic acid molecule of interest (e.g., a gene, cDNA molecule, or cDNA library) into a desired vector or into a host cell.

25

30 Using the information provided herein, such as the nucleotide sequences for the recombination site sequences described herein, an isolated nucleic acid molecule of the present invention encoding one or more recombination sites or portions thereof may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Preferred such

methods include PCR-based cloning methods, such as reverse transcriptase-PCR (RT-PCR) using primers such as those described herein and in the Examples below. Alternatively, vectors comprising the cassettes containing the recombination site sequences described herein are available commercially from  
5 Life Technologies, Inc. (Rockville, MD).

The invention is also directed to nucleic acid molecules comprising one or more of the recombination site sequences or portions thereof and one or more additional nucleotide sequences, which may encode functional or structural sites such as one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (which may be promoters, enhancers, repressors, and the like), one or more translational signals (e.g., secretion signal sequences), one or more origins of replication, one or more fusion partner peptides (particularly glutathione S-transferase (GST), hexahistidine (His<sub>6</sub>), and thioredoxin (Trx)), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more genes or portions of genes encoding a protein or polypeptide of interest, and one or more 5' polynucleotide extensions (particularly an extension of guanine residues ranging in length from about 1 to about 20, from about 2 to about 15, from about 3 to about 10, from about 4 to about 10, and most preferably an extension of 4 or 5 guanine residues at the 5' end of the recombination site nucleotide sequence. The one or more additional functional or structural sequences may or may not flank one or more of the recombination site sequences contained on the nucleic acid molecules of the  
10 invention.

15  
20  
25

In some nucleic acid molecules of the invention, the one or more nucleotide sequences encoding one or more additional functional or structural sites may be operably linked to the nucleotide sequence encoding the recombination site. For example, certain nucleic acid molecules of the invention may have a promoter sequence operably linked to a nucleotide sequence encoding a recombination site or portion thereof of the invention, such as a T7 promoter, a phage lambda PL  
30

promoter, an *E. coli lac*, *trp* or *tac* promoter, and other suitable promoters which will be familiar to the skilled artisan.

Nucleic acid molecules of the present invention, which may be isolated nucleic acid molecules, may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically, or in the form of DNA-RNA hybrids. The nucleic acid molecules of the invention may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. The nucleic acid molecules of the invention may also have a number of topologies, including linear, circular, coiled, or supercoiled.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells, and those DNA molecules purified (partially or substantially) from a solution whether produced by recombinant DNA or synthetic chemistry techniques. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention.

The present invention further relates to mutants, fragments, variants and derivatives of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of one or more recombination sites. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (see Lewin, B., ed., *Genes II*, John Wiley & Sons, New York (1985)). Non-naturally occurring variants may be produced using art-known mutagenesis techniques, such as those described hereinbelow.

Such variants include those produced by nucleotide substitutions, deletions or additions or portions thereof, or combinations thereof. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding

regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the encoded polypeptide(s) or portions thereof, and which also do not substantially alter the reactivities of the recombination site nucleic acid sequences in recombination reactions. Also especially preferred in this regard are conservative substitutions.

Particularly preferred mutants, fragments, variants, and derivatives of the nucleic acid molecules of the invention include, but are not limited to, insertions, deletions or substitutions of one or more nucleotide bases within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four wildtype lambda *att* sites, *attB*, *attP*, *attL* and *attR* (see U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, which describes the core region in further detail, and the disclosures of which are incorporated herein by reference in their entireties). Analogously, the core regions in *attB*1, *attP*1, *attL*1 and *attR*1 are identical to one another, as are the core regions in *attB*2, *attP*2, *attL*2 and *attR*2. Particularly preferred in this regard are nucleic acid molecules comprising insertions, deletions or substitutions of one or more nucleotides within the seven bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) that occurs within this 15 bp core region (GCTTTTTTATACTAA). Examples of such preferred mutants, fragments, variants and derivatives according to this aspect of the invention include, but are not limited to, nucleic acid molecules in which the thymine at position 1 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 2 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 3 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the adenine at position 4 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; in which the thymine at position 5 of the seven bp overlap region has been deleted or substituted with a

guanine, cytosine, or adenine; in which the adenine at position 6 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; and in which the cytosine at position 7 of the seven bp overlap region has been deleted or substituted with a guanine, thymine, or adenine; or any combination of one or more such deletions and/or substitutions within this seven bp overlap region. As described in detail in Example 21 herein, mutants of the nucleic acid molecules of the invention in which substitutions have been made within the first three positions of the seven bp overlap (TTTATAC) have been found in the present invention to strongly affect the specificity of recombination, mutant nucleic acid molecules in which substitutions have been made in the last four positions (TTTATAC) only partially alter recombination specificity, and mutant nucleic acid molecules comprising nucleotide substitutions outside of the seven bp overlap, but elsewhere within the 15 bp core region, do not affect specificity of recombination but do influence the efficiency of recombination.

Hence, in an additional aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that affect recombination specificity, particularly one or more nucleotide sequences that may correspond substantially to the seven base pair overlap within the 15 bp core region, having one or more mutations that affect recombination specificity. Particularly preferred such molecules may comprise a consensus sequence (described in detail in Example 21 herein) such as NNNATAC, wherein "N" refers to any nucleotide (*i.e.*, may be A, G, T/U or C), with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

In a related aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that enhance recombination efficiency, particularly one or more nucleotide sequences that may correspond substantially to the core region and having one or more mutations that enhance recombination efficiency. By sequences or mutations that "enhance recombination efficiency" is meant a sequence or mutation in a recombination site, preferably in the core region (*e.g.*, the 15 bp core region of *att* recombination sites), that results in an increase in cloning efficiency (typically

measured by determining successful cloning of a test sequence, e.g., by determining CFU/ml for a given cloning mixture) when recombining molecules comprising the mutated sequence or core region as compared to molecules that do not comprise the mutated sequence or core region (e.g., those comprising a wildtype recombination site core region sequence). More specifically, whether or not a given sequence or mutation enhances recombination efficiency may be determined using the sequence or mutation in recombinational cloning as described herein, and determining whether the sequence or mutation provides enhanced recombinational cloning efficiency when compared to a non-mutated (e.g., wildtype) sequence. Methods of determining preferred cloning efficiency-enhancing mutations for a number of recombination sites, particularly for *att* recombination sites, are described herein, for example in Examples 22-25. Examples of preferred such mutant recombination sites include but are not limited to the *attL* consensus core sequence of caactnnntnnnannaagttg (wherein "n" represents any nucleotide), for example the *attL5* sequence agcctgctttattatactaagttggcatta and the *attL6* sequence agcctgcttttatattaagttggcatta; the *attB1.6* sequence ggggacaacttgtacaaaaaagttggct; the *attB2.2* sequence ggggacaacttgtacaagaaagctgggt; and the *attB2.10* sequence ggggacaacttgtacaagaaagtgggt. Those of skill in the art will appreciate that, in addition to the core region, other portions of the *att* site may affect the efficiency of recombination. There are five so-called arm binding sites for the integrase protein in the bacteriophage lambda *attP* site, two in *attR* (P1 and P2), and three in *attL* (P'1, P'2 and P'3). Compared to the core binding sites, the integrase protein binds to arm sites with high affinity and interacts with core and arm sites through two different domains of the protein. As with the core binding site a consensus sequence for the arm binding site consisting of C/AAGTCACTAT has been inferred from sequence comparison of the five arm binding sites and seven non-*att* sites (Ross and Landy, *Proc. Natl. Acad. Sci. USA* 79:7724-7728 (1982)). Each arm site has been mutated and tested for its effect in the excision and integration reactions (Numrych *et al.*, *Nucl. Acids Res.* 18:3953 (1990)). Hence, specific sites are utilized in each reaction in different ways, namely, the P1 and P'3

sites are essential for the integration reaction whereas the other three sites are dispensable to the integration reaction to varying degrees. Similarly, the P2, P'1 and P'2 sites are most important for the excision reaction, whereas P1 and P'3 are completely dispensable. Interestingly, when P2 is mutated the integration reaction occurs more efficiently than with the wild type attP site. Similarly, when P1 and P'3 are mutated the excision reaction occurs more efficiently. The stimulatory effect of mutating integrase arm binding sites can be explained by removing sites that compete or inhibit a specific recombination pathway or that function in a reaction that converts products back to starting substrates. In fact there is evidence for an XIS-independent LR reaction (Abremski and Gottesman, *J. Mol. Biol.* 153:67-78 (1981)). Thus, in addition to modifications in the core region of the att site, the present invention contemplates the use of att sites containing one or more modifications in the integrase arm-type binding sites. In some preferred embodiments, one or more mutations may be introduced into one or more of the P1, P'1, P2, P'2 and P'3 sites. In some preferred embodiments, multiple mutations may be introduced into one or more of these sites. Preferred such mutations include those which increase the recombination *in vitro*. For example, in some embodiments mutations may be introduced into the arm-type binding sites such that integrative recombination, corresponding to the BP reaction, is enhanced. In other embodiments, mutations may be introduced into the arm-type binding sites such that excisive recombination, corresponding to the LR reaction, is enhanced. Of course, based on the guidance contained herein, particularly in the construction and evaluation of effects of mutated recombination sites upon recombinational specificity and efficiency, analogous mutated or engineered sequences may be produced for other recombination sites described herein (including but not limited to *lox*, FRT, and the like) and used in accordance with the invention. For example, much like the mutagenesis strategy used to select core binding sites that enhance recombination efficiency, similar strategies can be employed to select changes in the arms of attP, attL and attR, and in analogous sequences in other recombination sites such as *lox*, FRT and the like, that enhance recombination efficiency. Hence, the construction and evaluation of such mutants is well within the abilities of those of ordinary skill in the art without undue experimentation.

One suitable methodology for preparing and evaluating such mutations is found in Numrych, *et al.*, (1990) *Nucleic Acids Research* 18(13): 3953-3959.

Other mutant sequences and nucleic acid molecules that may be suitable to enhance recombination efficiency will be apparent from the description herein, or may be easily determined by one of ordinary skill using only routine experimentation in molecular biology in view of the description herein and information that is readily available in the art

Since the genetic code is well known in the art, it is also routine for one of ordinary skill in the art to produce degenerate variants of the nucleic acid molecules described herein without undue experimentation. Hence, nucleic acid molecules comprising degenerate variants of nucleic acid sequences encoding the recombination sites described herein are also encompassed within the scope of the invention.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 50% identical, at least 60% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to the nucleotide sequences of the seven bp overlap region within the 15 bp core region of the recombination sites described herein, or the nucleotide sequences of *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* as set forth in Figure 9 (or portions thereof), or a nucleotide sequence complementary to any of these nucleotide sequences, or fragments, variants, mutants, and derivatives thereof.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a particular recombination site or portion thereof is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations (*e.g.*, insertions, substitutions, or deletions) per each 100 nucleotides of the reference nucleotide sequence encoding the recombination site. For example, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference *attB1* nucleotide sequence, up to 5% of the nucleotides in the *attB1* reference sequence may be

5 deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the *attB1* reference sequence may be inserted into the *attB1* reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

10 As a practical matter, whether any particular nucleic acid molecule is at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, a given recombination site nucleotide sequence or portion thereof can be determined conventionally using known computer programs such as DNAsis software (Hitachi Software, San Bruno, California) for initial sequence alignment followed by ESEE version 3.0 DNA/protein sequence software (cabot@trog.mbb.sfu.ca) for multiple sequence alignments. Alternatively, such 15 determinations may be accomplished using the BESTFIT program (Wisconsin Sequence Analysis Package, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711), which employs a local homology algorithm (Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981)) to find the best segment of homology between two sequences. When 20 using DNAsis, ESEE, BESTFIT or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number 25 of nucleotides in the reference sequence are allowed.

30 The present invention is directed to nucleic acid molecules at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences as set forth in Figure 9, or to the nucleotide sequence of the deposited clones, irrespective of whether they encode particular functional polypeptides. This is because even where a particular nucleic acid molecule does not encode a particular functional polypeptide, one of skill in the art would still know how to use the nucleic acid

molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer.

Mutations can also be introduced into the recombination site nucleotide sequences for enhancing site specific recombination or altering the specificities of the reactants, etc. Such mutations include, but are not limited to: recombination sites without translation stop codons that allow fusion proteins to be encoded; recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively with their homologous partners allowing multiple reactions to be contemplated; and mutations that prevent hairpin formation of recombination sites. Which particular reactions take place can be specified by which particular partners are present in the reaction mixture.

There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel, F.M. *et al.*, *Current Protocols in Molecular Biology*, Wiley Interscience, New York (1989-1996). Mutations can be designed into oligonucleotides, which can be used to modify existing cloned sequences, or in amplification reactions. Random mutagenesis can also be employed if appropriate selection methods are available to isolate the desired mutant DNA or RNA. The presence of the desired mutations can be confirmed by sequencing the nucleic acid by well known methods.

The following non-limiting methods can be used to modify or mutate a given nucleic acid molecule encoding a particular recombination site to provide mutated sites that can be used in the present invention:

1. By recombination of two parental DNA sequences by site-specific (e.g. attL and attR to give attP) or other (e.g. homologous) recombination mechanisms where the parental DNA segments contain one or more base alterations resulting in the final mutated nucleic acid molecule;
2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired nucleic acid molecule;
3. By mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired nucleic acid molecule;

4. By reverse transcription of an RNA encoding the desired core sequence; and
5. By *de novo* synthesis (chemical synthesis) of a sequence having the desired base changes, or random base changes followed by sequencing or functional analysis according to methods that are routine in the art.

The functionality of the mutant recombination sites can be demonstrated in ways that depend on the particular characteristic that is desired. For example, the lack of translation stop codons in a recombination site can be demonstrated by expressing the appropriate fusion proteins. Specificity of recombination between homologous partners can be demonstrated by introducing the appropriate molecules into *in vitro* reactions, and assaying for recombination products as described herein or known in the art. Other desired mutations in recombination sites might include the presence or absence of restriction sites, translation or transcription start signals, protein binding sites, particular coding sequences, and other known functionalities of nucleic acid base sequences. Genetic selection schemes for particular functional attributes in the recombination sites can be used according to known method steps. For example, the modification of sites to provide (from a pair of sites that do not interact) partners that do interact could be achieved by requiring deletion, via recombination between the sites, of a DNA sequence encoding a toxic substance. Similarly, selection for sites that remove translation stop sequences, the presence or absence of protein binding sites, etc., can be easily devised by those skilled in the art.

Accordingly, the present invention also provides a nucleic acid molecule, comprising at least one DNA segment having at least one, and preferably at least two, engineered recombination site nucleotide sequences of the invention flanking a selectable marker and/or a desired DNA segment, wherein at least one of said recombination site nucleotide sequences has at least one engineered mutation that enhances recombination *in vitro* in the formation of a Cointegrate DNA or a Product DNA. Such engineered mutations may be in the core sequence of the recombination site nucleotide sequence of the invention; *see* U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed

October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties.

While in the preferred embodiment the recombination sites differ in sequence and do not interact with each other, it is recognized that sites comprising the same sequence, which may interact with each other, can be manipulated or engineered to inhibit recombination with each other. Such conceptions are considered and incorporated herein. For example, a protein binding site (*e.g.*, an antibody-binding site, a histone-binding site, an enzyme-binding site, or a binding site for any nucleic acid molecule-binding protein) can be engineered adjacent to one of the sites. In the presence of the protein that recognizes the engineered site, the recombinase fails to access the site and another recombination site in the nucleic acid molecule is therefore used preferentially. In the cointegrate this site can no longer react since it has been changed, *e.g.*, from attB to attL. During or upon resolution of the cointegrate, the protein can be inactivated (*e.g.*, by antibody, heat or a change of buffer) and the second site can undergo recombination.

The nucleic acid molecules of the invention can have at least one mutation that confers at least one enhancement of said recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (iii) relieving the requirement for host factors; (iv) increasing the efficiency of said Cointegrate DNA or Product DNA formation; (v) increasing the specificity of said Cointegrate DNA or Product DNA formation; and (vi) adding or deleting protein binding sites.

In other embodiments, the nucleic acid molecules of the invention may be PCR primer molecules, which comprise one or more of the recombination site sequences described herein or portions thereof, particularly those shown in Figure 9 (or sequences complementary to those shown in Figure 9), or mutants, fragments, variants or derivatives thereof, attached at the 3' end to a target-specific template sequence which specifically interacts with a target nucleic acid molecule which is to be amplified. Primer molecules according to this aspect of the invention may further comprise one or more, (*e.g.*, 1, 2, 3, 4, 5, 10, 20, 25, 50, 100, 500, 1000, or more) additional bases at their 5' ends, and preferably comprise one or more (particularly four or five) additional bases, which are preferably

guanines, at their 5' ends, to increase the efficiency of the amplification products incorporating the primer molecules in the recombinational cloning system of the invention. Such nucleic acid molecules and primers are described in detail in the examples herein, particularly in Examples 22-25.

5        Certain primers of the invention may comprise one or more nucleotide deletions in the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* sequences as set forth in Figure 9. In one such aspect, for example, *attB2* primers may be constructed in which one or more of the first four nucleotides at the 5' end of the *attB2* sequence shown in Figure 9 have been deleted. Primers according to this aspect of the invention may therefore have the sequence:

10      (*attB2(-1)*): CCCAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2(-2)*): CCAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n

(*attB2(-3)*): CAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n

(*attB2(-4)*): AGCTTCTTGTACAAAGTGGTnnnnnnnnnnnnnn . . . n,

15      wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

20      The primer nucleic acid molecules according to this aspect of the invention may be produced synthetically by attaching the recombination site sequences depicted in Figure 9, or portions thereof, to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art. Alternatively, additional primer nucleic acid molecules of the invention may be produced synthetically by adding one or more nucleotide bases, which preferably correspond to one or more, preferably five or more, and more preferably six or more, contiguous nucleotides of the *att* nucleotide sequences described herein (see, e.g., Example 20 herein; see also U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties), to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art, to provide

primers having the specific nucleotide sequences described herein. As noted above, primer nucleic acid molecules according to this aspect of the invention may also optionally comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four or five guanines at their 5' ends. In one particularly preferred such aspect, the primer nucleic acid molecules of the invention may comprise one or more, preferably five or more, more preferably six or more, still more preferably 6-18 or 6-25, and most preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25, contiguous nucleotides or bp of the *attB1* or *attB2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. Primer nucleic acid molecules according to this aspect of the invention include, but are not limited to, *attB1*- and *attB2*-derived primer nucleic acid molecules having the following nucleotide sequences:

ACAAGTTGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n

ACCACTTGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n

TGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n

TGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n

ACAAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n

ACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n

AAAAAGCAGGCT-nnnnnnnnnnnnn . . . n

AGAAAGCTGGGT-nnnnnnnnnnnnn . . . n

AAAAGCAGGCT-nnnnnnnnnnnnn . . . n

GAAAGCTGGGT-nnnnnnnnnnnnn . . . n

AAAGCAGGCT-nnnnnnnnnnnnn . . . n

AAAGCTGGGT-nnnnnnnnnnnnn . . . n

AAGCAGGCT-nnnnnnnnnnnnn . . . n

AAGCTGGGT-nnnnnnnnnnnnn . . . n

AGCAGGCT-nnnnnnnnnnnnn . . . n

AGCTGGGT-nnnnnnnnnnnnn . . . n

GCAGGCT-nnnnnnnnnnnnn . . . n

GCTGGGT-nnnnnnnnnnnnn . . . n

CAGGCT-nnnnnnnnnnnn . . . n

CTGGGT-nnnnnnnnnnnn . . . n,

wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

Of course, it will be apparent to one of ordinary skill from the teachings contained herein that additional primer nucleic acid molecules analogous to those specifically described herein may be produced using one or more, preferably five or more, more preferably six or more, still more preferably ten or more, 15 or more, 20 or more, 25 or more, 30 or more, etc. (through to and including all) of the contiguous nucleotides or bp of the *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. As noted above, such primer nucleic acid molecules may optionally further comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four guanines at their 5' ends. Other primer molecules comprising the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* sequences depicted in Figure 9, or portions thereof, may be made by one of ordinary skill without resorting to undue experimentation in accordance with the guidance provided herein.

The primers of the invention described herein are useful in producing PCR fragments having a nucleic acid molecule of interest flanked at each end by a recombination site sequence (as described in detail below in Example 9), for use in cloning of PCR-amplified DNA fragments using the recombination system of the invention (as described in detail below in Examples 8, 19 and 21-25).

### *Vectors*

The invention also relates to vectors comprising one or more of the nucleic acid molecules of the invention, as described herein. In accordance with the invention, any vector may be used to construct the vectors of the invention. In

particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with the invention be engineered to include one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), or mutants, fragments, or derivatives thereof, for use in the methods of the invention. Such vectors may be obtained from, for example, 5 Vector Laboratories Inc., InVitrogen, Promega, Novagen, New England Biolabs, Clontech, Roche, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc., and Research Genetics. Such 10 vectors may then for example be used for cloning or subcloning nucleic acid molecules of interest. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, Expression Vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large inserts and the like.

15 Other vectors of interest include viral origin vectors (M13 vectors, bacterial phage  $\lambda$  vectors, bacteriophage P1 vectors, adenovirus vectors, herpesvirus vectors, retrovirus vectors, phage display vectors, combinatorial library vectors), high, low, and adjustable copy number vectors, vectors which have compatible 20 replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8).

25 Particular vectors of interest include prokaryotic Expression Vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Inc.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZ218, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Life Technologies, Inc.) and variants and derivatives thereof. Destination Vectors can also be made from eukaryotic Expression 30 Vectors such as pFastBac, pFastBac HT, pFastBac DUAL, pSFV, and pTet-Splice (Life Technologies, Inc.), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A,

B, and C, pVL1392, pBsueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Inc.) and variants or derivatives thereof.

Other vectors of particular interest include pUC18, pUC19, pBlueScript, pSPORT, cosmids, phagemids, YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), MACs (mammalian artificial chromosomes), pQE70, pQE60, pQE9 (Qiagen), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (InVitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Life Technologies, Inc.) and variants or derivatives thereof.

Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ $\alpha$ , pGAPZ, pGAPZ $\alpha$ , pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZErO1.1, pZErO-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe, SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen;  $\lambda$ ExCell,  $\lambda$ gt11, pTrc99A, pKK223-3, pGEX-1 $\lambda$ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZ218, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTAG, pET-32LIC, pET-30LIC, pBAC-2cpLIC, pBACgus-2cpLIC, pT7Blue-2LIC, pT7Blue-2,  $\lambda$ SCREEN-1,  $\lambda$ BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta - Hyg, and Selecta Vecta - Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1,

5 pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP,  
pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-Basic,  
pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p $\beta$ gal-Basic,  
p $\beta$ gal-Control, p $\beta$ gal-Promoter, p $\beta$ gal-Enhancer, pCMV $\beta$ , pTet-Off, pTet-On,  
pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX,  
pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo,  
pYEX 4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6,  
pTriplEx,  $\lambda$ gt10,  $\lambda$ gt11, pWE15, and  $\lambda$ TriplEx from Clontech; Lambda ZAP II,  
10 pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4,  
pBD-GAL4 Cam, pSurfscript, Lambda FIX II, Lambda DASH, Lambda EMBL3,  
Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script  
Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n,  
15 pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLacI,  
pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo  
Poly A, pOG44, pOG45, pFRT $\beta$ GAL, pNEO $\beta$ GAL, pRS403, pRS404, pRS405,  
pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

20 Two-hybrid and reverse two-hybrid vectors of particular interest include  
pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pACt, pACT2,  
pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4,  
pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202,  
pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

25 Yeast Expression Vectors of particular interest include pESP-1, pESP-2,  
pESC-His, pESC-Trp, pESC-URA, pESC-Leu (Stratagene), pRS401, pRS402,  
pRS411, pRS412, pRS421, pRS422, and variants or derivatives thereof.

According to the invention, the vectors comprising one or more nucleic acid  
30 molecules encoding one or more recombination sites, or mutants, variants,  
fragments, or derivatives thereof, may be produced by one of ordinary skill in the  
art without resorting to undue experimentation using standard molecular biology  
methods. For example, the vectors of the invention may be produced by  
introducing one or more of the nucleic acid molecules encoding one or more  
recombination sites (or mutants, fragments, variants or derivatives thereof) into  
one or more of the vectors described herein, according to the methods described,

for example, in Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). In a related aspect of the invention, the vectors may be engineered to contain, in addition to one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), one or more additional physical or functional nucleotide sequences, such as those encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (e.g., one or more promoters, enhancers, or repressors), one or more selection markers or modules, one or more genes or portions of genes encoding a protein or polypeptide of interest, one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (e.g., GST, His<sub>6</sub> or thioredoxin), one or more origins of replication, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). According to this aspect of the invention, the one or more recombination site nucleotide sequences (or portions thereof) may optionally be operably linked to the one or more additional physical or functional nucleotide sequences described herein.

Preferred vectors according to this aspect of the invention include, but are not limited to: pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), pENTR3C (Figures 12A and 12B), pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), pENTR6 (Figures 15A and 15B), pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), pENTR9 (Figures 18A and 18B), pENTR10 (Figures 19A and 19B), pENTR11 (Figures 20A and 20B), pDEST1 (Figures 21A-D), pDEST2 (Figure 22A-D), pDEST3 (Figure 23A-D), pDEST4 (Figure 24A-D), pDEST5 (Figure 25A-D), pDEST6 (Figure 26A-D), pDEST7 (Figure 27A-C), pDEST8 (Figure 28A-D), pDEST9 (Figure 29A-E), pDEST10 (Figure 30A-D), pDEST11 (Figure 31A-D), pDEST12.2 (also known as pDEST12) (Figure 32A-D), pDEST13 (Figure 33A-C), pDEST14 (Figure 34A-D), pDEST15 (Figure 35A-D), pDEST16 (Figure 36A-D), pDEST17 (Figure 37A-D), pDEST18 (Figure 38A-D), pDEST19 (Figure 39A-D), pDEST20 (Figure 40A-D), pDEST21 (Figure 41A-E), pDEST22 (Figure 42A-D), pDEST23 (Figure 43A-D), pDEST24 (Figure 44A-D), pDEST25 (Figure 45A-D), pDEST26 (Figure 46A-D), pDEST27 (Figure 47A-D), pEXP501 (also known

as pCMVSPORT6) (Figure 48A-B), pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector) (Figure 49), pDONR202 (Figure 50), pDONR203 (also known as pEZ15812) (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector) (Figure 54), pMAB58 (Figure 87), pMAB62 (Figure 88), pDEST28 (Figure 90), pDEST29 (Figure 91), pDEST30 (Figure 92), pDEST31 (Figure 93), pDEST32 (Figure 94), pDEST33 (Figure 95), pDEST34 (Figure 96), pDONR207 (Figure 97), pMAB85 (Figure 98), pMAB86 (Figure 99), and fragments, mutants, variants, and derivatives thereof. However, it will be understood by one of ordinary skill that the present invention also encompasses other vectors not specifically designated herein, which comprise one or more of the isolated nucleic acid molecules of the invention encoding one or more recombination sites or portions thereof (or mutants, fragments, variants or derivatives thereof), and which may further comprise one or more additional physical or functional nucleotide sequences described herein which may optionally be operably linked to the one or more nucleic acid molecules encoding one or more recombination sites or portions thereof. Such additional vectors may be produced by one of ordinary skill according to the guidance provided in the present specification.

20

### *Polymerases*

Preferred polypeptides having reverse transcriptase activity (*i.e.*, those polypeptides able to catalyze the synthesis of a DNA molecule from an RNA template) for use in accordance with the present invention include, but are not limited to Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, Myeloblastosis Associated Virus (MAV) reverse transcriptase, Human Immunodeficiency Virus (HTV) reverse transcriptase, retroviral reverse transcriptase, retrotransposon reverse transcriptase, hepatitis B reverse transcriptase, cauliflower mosaic virus reverse transcriptase and bacterial reverse transcriptase. Particularly preferred are those polypeptides having reverse

transcriptase activity that are also substantially reduced in RNase H activity (*i.e.*, "RNase H" polypeptides). By a polypeptide that is "substantially reduced in RNase H activity" is meant that the polypeptide has less than about 20%, more preferably less than about 15%, 10% or 5%, and most preferably less than about 2%, of the RNase H activity of a wildtype or RNase H<sup>-</sup> enzyme such as wildtype M-MLV reverse transcriptase. The RNase H activity may be determined by a variety of assays, such as those described, for example, in U.S. Patent No. 5,244,797, in Kotewicz, M.L. *et al.*, *Nucl. Acids Res.* 16:265 (1988) and in Gerard, G.F., *et al.*, *FOCUS* 14(5):91 (1992), the disclosures of all of which are fully incorporated herein by reference. Suitable RNase H<sup>-</sup> polypeptides for use in the present invention include, but are not limited to, M-MLV H<sup>-</sup> reverse transcriptase, RSV H<sup>-</sup> reverse transcriptase, AMV H<sup>-</sup> reverse transcriptase, RAV H<sup>-</sup> reverse transcriptase, MAV H<sup>-</sup> reverse transcriptase, HIV H<sup>-</sup> reverse transcriptase, THERMOSCRIPT™ reverse transcriptase and THERMOSCRIPT™ II reverse transcriptase, and SUPERSCRIPT™ I reverse transcriptase and SUPERSCRIPT™ II reverse transcriptase, which are obtainable, for example, from Life Technologies, Inc. (Rockville, Maryland). See generally published PCT application WO 98/47912.

Other polypeptides having nucleic acid polymerase activity suitable for use in the present methods include thermophilic DNA polymerases such as DNA polymerase I, DNA polymerase III, Klenow fragment, T7 polymerase, and T5 polymerase, and thermostable DNA polymerases including, but not limited to, *Thermus thermophilus* (*Tth*) DNA polymerase, *Thermus aquaticus* (*Taq*) DNA polymerase, *Thermotoga neopolitana* (*Tne*) DNA polymerase, *Thermotoga maritima* (*Tma*) DNA polymerase, *Thermococcus litoralis* (*Tli* or VENT®) DNA polymerase, *Pyrococcus furiosus* (*Pfu*) DNA polymerase, *Pyrococcus* species GB-D (or DEEPVENT®) DNA polymerase, *Pyrococcus woosii* (*Pwo*) DNA polymerase, *Bacillus stearothermophilus* (*Bst*) DNA polymerase, *Sulfolobus acidocaldarius* (*Sac*) DNA polymerase, *Thermoplasma acidophilum* (*Tac*) DNA polymerase, *Thermus flavus* (*Tfl/Tub*) DNA polymerase, *Thermus ruber* (*Tru*) DNA polymerase, *Thermus brockianus* (DYNAZYME®) DNA polymerase, *Methanobacterium thermoautotrophicum* (*Mth*) DNA polymerase, and mutants,

variants and derivatives thereof. Such polypeptides are available commercially, for example from Life Technologies, Inc. (Rockville, MD), New Englan BioLabs (Beverly, MA), and Sigma/Aldrich (St. Louis, MO).

5

### *Host Cells*

10

The invention also relates to host cells comprising one or more of the nucleic acid molecules or vectors of the invention, particularly those nucleic acid molecules and vectors described in detail herein. Representative host cells that may be used according to this aspect of the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include *Escherichia* spp. cells (particularly *E. coli* cells and most particularly *E. coli* strains DH10B, Stbl2, DH5 $\alpha$ , DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety), *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly CHO, COS, VERO, BHK and human cells). Preferred yeast host cells include *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example from Life Technologies, Inc. (Rockville, Maryland), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois).

20

25

30

Methods for introducing the nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the nucleic acid molecules and/or vectors of the invention, will be

5

10

15

20

30

familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other the nucleic acid molecules and/or vectors. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as *E. coli*. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., *et al.*, *Molecular Cloning, a Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16.55 (1989), Watson, J.D., *et al.*, *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., pp. 213-234 (1992), and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures.

### ***Polypeptides***

In another aspect, the invention relates to polypeptides encoded by the nucleic acid molecules of the invention (including polypeptides and amino acid sequences encoded by all possible reading frames of the nucleic acid molecules of the invention), and to methods of producing such polypeptides. Polypeptides of the present invention include purified or isolated natural products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, insect, mammalian, avian and higher plant cells.

The polypeptides of the invention may be produced by synthetic organic chemistry, and are preferably produced by standard recombinant methods, employing one or more of the host cells of the invention comprising the vectors or isolated nucleic acid molecules of the invention. According to the invention, 5 polypeptides are produced by cultivating the host cells of the invention (which comprise one or more of the nucleic acid molecules of the invention, preferably contained within an Expression Vector) under conditions favoring the expression of the nucleotide sequence contained on the nucleic acid molecule of the invention, such that the polypeptide encoded by the nucleic acid molecule of the invention 10 is produced by the host cell. As used herein, "conditions favoring the expression of the nucleotide sequence" or "conditions favoring the production of a polypeptide" include optimal physical (e.g., temperature, humidity, etc.) and nutritional (e.g., culture medium, ionic) conditions required for production of a recombinant polypeptide by a given host cell. Such optimal conditions for a 15 variety of host cells, including prokaryotic (bacterial), mammalian, insect, yeast, and plant cells will be familiar to one of ordinary skill in the art, and may be found, for example, in Sambrook, J., et al., *Molecular Cloning, A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, (1989), Watson, J.D., et al., *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers 20 (1987).

In some aspects, it may be desirable to isolate or purify the polypeptides of the invention (e.g., for production of antibodies as described below), resulting in the production of the polypeptides of the invention in isolated form. The polypeptides 25 of the invention can be recovered and purified from recombinant cell cultures by well-known methods of protein purification that are routine in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. For example, His6 or GST fusion tags on polypeptides made by the methods of the invention may be isolated using 30 appropriate affinity chromatography matrices which bind polypeptides bearing

His<sub>6</sub> or GST tags, as will be familiar to one of ordinary skill in the art. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Isolated polypeptides of the invention include those comprising the amino acid sequences encoded by one or more of the reading frames of the polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or nucleotide sequences complementary thereto), or fragments, variants, mutants and derivatives thereof; the complete amino acid sequences encoded by the polynucleotides contained in the deposited clones described herein; the amino acid sequences encoded by polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences of the invention as set forth in Figure 9 (or a nucleotide sequence complementary thereto); or a peptide or polypeptide comprising a portion or a fragment of the above polypeptides. The invention also relates to additional polypeptides having one or more additional amino acids linked (typically by peptidyl bonds to form a nascent polypeptide) to the polypeptides encoded by the recombination site nucleotide sequences or the deposited clones. Such additional amino acid residues may comprise one or more functional peptide sequences, for example one or more fusion partner peptides (e.g., GST, His<sub>6</sub>, Trx, etc.) and the like.

As used herein, the terms "protein," "peptide," "oligopeptide" and "polypeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of two or more amino acids, preferably five or more amino acids, or more preferably ten

or more amino acids, coupled by (a) peptidyl linkage(s), unless otherwise defined in the specific contexts below. As is commonly recognized in the art, all polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus.

It will be recognized by those of ordinary skill in the art that some amino acid sequences of the polypeptides of the invention can be varied without significant effect on the structure or function of the polypeptides. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine structure and activity. In general, it is possible to replace residues which form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the polypeptide.

Thus, the invention further includes variants of the polypeptides of the invention, including allelic variants, which show substantial structural homology to the polypeptides described herein, or which include specific regions of these polypeptides such as the portions discussed below. Such mutants may include deletions, insertions, inversions, repeats, and type substitutions (for example, substituting one hydrophilic residue for another, but not strongly hydrophilic for strongly hydrophobic as a rule). Small changes or such "neutral" or "conservative" amino acid substitutions will generally have little effect on activity.

Typical conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxylated residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amidated residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr.

Thus, the fragment, derivative or analog of the polypeptides of the invention, such as those comprising peptides encoded by the recombination site nucleotide sequences described herein, may be (i) one in which one or more of the amino acid residues are substituted with a conservative or non-conservative amino acid residue (preferably a conservative amino acid residue), and such substituted amino acid residue may be encoded by the genetic code or may be an amino acid (e.g.,

desmosine, citrulline, ornithine, etc.) that is not encoded by the genetic code; (ii) one in which one or more of the amino acid residues includes a substituent group (e.g., a phosphate, hydroxyl, sulfate or other group) in addition to the normal "R" group of the amino acid; (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which additional amino acids are fused to the mature polypeptide, such as an immunoglobulin Fc region peptide, a leader or secretory sequence, a sequence which is employed for purification of the mature polypeptide (such as GST) or a proprotein sequence. Such fragments, derivatives and analogs are intended to be encompassed by the present invention, and are within the scope of those skilled in the art from the teachings herein and the state of the art at the time of invention.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Recombinantly produced versions of the polypeptides of the invention can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). As used herein, the term "substantially purified" means a preparation of an individual polypeptide of the invention wherein at least 50%, preferably at least 60%, 70%, or 75% and more preferably at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% (by mass) of contaminating proteins (*i.e.*, those that are not the individual polypeptides described herein or fragments, variants, mutants or derivatives thereof) have been removed from the preparation.

The polypeptides of the present invention include those which are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the polypeptides described herein. For example, preferred *attB1*-containing polypeptides of the invention include those that are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical,

to the polypeptide(s) encoded by the three reading frames of a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto), to a polypeptide encoded by a polynucleotide contained in the deposited cDNA clones described herein, or to a polypeptide encoded by a polynucleotide hybridizing under stringent conditions to a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Analogous polypeptides may be prepared that are at least about 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* polypeptides of the invention as depicted in Figure 9. The present polypeptides also include portions or fragments of the above-described polypeptides with at least 5, 10, 15, 20, or 25 amino acids.

By a polypeptide having an amino acid sequence at least, for example, 65% "identical" to a reference amino acid sequence of a given polypeptide of the invention is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to 35 amino acid alterations per each 100 amino acids of the reference amino acid sequence of a given polypeptide of the invention. In other words, to obtain a polypeptide having an amino acid sequence at least 65% identical to a reference amino acid sequence, up to 35% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 35% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino (N-) or carboxy (C-) terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence. As a practical matter, whether a given amino acid sequence is, for example, at least 65% identical to the amino acid sequence of a given polypeptide of the invention can be determined

conventionally using known computer programs such as those described above for nucleic acid sequence identity determinations, or more preferably using the CLUSTAL W program (Thompson, J.D., *et al.*, *Nucleic Acids Res.* 22:4673-4680 (1994)).

5       The polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. In addition, as described in detail below, the polypeptides of the present invention can be used to raise polyclonal and monoclonal antibodies which are useful in a variety of assays for detecting protein expression, localization, detection of interactions with other molecules, or 10 for the isolation of a polypeptide (including a fusion polypeptide) of the invention.

15      In another aspect, the present invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention, which may be used to raise antibodies, particularly monoclonal antibodies, that bind specifically to a one or more of the polypeptides of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. 20 On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (*see, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998- 4002 (1983)).

25      As to the selection of peptides or polypeptides bearing an antigenic epitope (*i.e.*, that contain a region of a protein molecule to which an antibody can bind), it is well-known in the art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein (*see, e.g.*, Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are not confined to the immunodominant regions of intact proteins (*i.e.*, immunogenic epitopes) or to the amino or carboxy

termini. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer peptides, especially those containing proline residues, usually are effective (Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)).

5       Epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least five, more preferably at least seven or more amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a given polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (*i.e.*, the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); sequences containing proline residues are particularly preferred.

10      Non-limiting examples of epitope-bearing polypeptides or peptides that can be used to generate antibodies specific for the polypeptides of the invention include certain epitope-bearing regions of the polypeptides comprising amino acid sequences encoded by polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or a nucleotide sequence complementary thereto); the complete amino acid sequences encoded by the three reading frames of the polynucleotides contained in the deposited clones described herein; and the amino acid sequences encoded by all reading frames of polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences (or portions thereof) of the invention as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Other epitope-bearing polypeptides or peptides that may be used to generate antibodies specific for the polypeptides

of the invention will be apparent to one of ordinary skill in the art based on the primary amino acid sequences of the polypeptides of the invention described herein, via the construction of Kyte-Doolittle hydrophilicity and Jameson-Wolf antigenic index plots of the polypeptides of the invention using, for example, PROTEAN computer software (DNASTAR, Inc.; Madison, Wisconsin).

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis (see, e.g., U.S. Patent No. 4,631,211 and Houghten, R. A., *Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985), both of which are incorporated by reference herein in their entireties).

As one of skill in the art will appreciate, the polypeptides of the present invention and epitope-bearing fragments thereof may be immobilized onto a solid support, by techniques that are well-known and routine in the art. By "solid support" is intended any solid support to which a peptide can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Linkage of the peptide of the invention to a solid support can be accomplished by attaching one or both ends of the peptide to the support. Attachment may also be made at one or more internal sites in the peptide. Multiple attachments (both internal and at the ends of the peptide) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulphhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments to the support, addition of an affinity tag sequence to the peptide can be used such as GST (Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., et al., *J. Chromatog.* 411:77 (1987)), or biotin. Such affinity tags

5 may be used for the reversible attachment of the peptide to the support. Such immobilized polypeptides or fragments may be useful, for example, in isolating antibodies directed against one or more of the polypeptides of the invention, or other proteins or peptides that recognize other proteins or peptides that bind to one or more of the polypeptides of the invention, as described below.

10 As one of skill in the art will also appreciate, the polypeptides of the present invention and the epitope-bearing fragments thereof described herein can be combined with one or more fusion partner proteins or peptides, or portions thereof, including but not limited to GST, His<sub>6</sub>, Trx, and portions of the constant domain of immunoglobulins (Ig), resulting in chimeric or fusion polypeptides. These fusion polypeptides facilitate purification of the polypeptides of the invention (EP 0 394 827; Traunecker *et al.*, *Nature* 331:84- 86 (1988)) for use in analytical or diagnostic (including high-throughput) format.

15 **Antibodies**

20 In another aspect, the invention relates to antibodies that recognize and bind to the polypeptides (or epitope-bearing fragments thereof) or nucleic acid molecules (or portions thereof) of the invention. In a related aspect, the invention relates to antibodies that recognize and bind to one or more polypeptides encoded by all reading frames of one or more recombination site nucleic acid sequences or portions thereof, or to one or more nucleic acid molecules comprising one or more recombination site nucleic acid sequences or portions thereof, including but not limited to *att* sites (including *att*B1, *att*B2, *att*P1, *att*P2, *att*L1, *att*L2, *att*R1, *att*R2 and the like), *lox* sites (e.g., *lox*P, *lox*P511, and the like), FRT, and the like, or mutants, fragments, variants and derivatives thereof. See generally U.S. Patent No. 5,888,732, which is incorporated herein by reference in its entirety. The antibodies of the present invention may be polyclonal or monoclonal, and may be prepared by any of a variety of methods and in a variety of species according to methods that are well-known in the art. See, for instance, U.S. Patent No. 5,587,287; Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983); Wilson *et al.*, *Cell* 37: 767 (1984); and Bittle, F.J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985). Antibodies specific for any of the polypeptides or nucleic acid molecules described

herein, such as antibodies specifically binding to one or more of the polypeptides encoded by the recombination site nucleotide sequences, or one or more nucleic acid molecules, described herein or contained in the deposited clones, antibodies against fusion polypeptides (*e.g.*, binding to fusion polypeptides between one or more of the fusion partner proteins and one or more of the recombination site polypeptides of the invention, as described herein), and the like, can be raised against the intact polypeptides or polynucleotides of the invention or one or more antigenic polypeptide fragments thereof.

As used herein, the term "antibody" (Ab) may be used interchangeably with the terms "polyclonal antibody" or "monoclonal antibody" (mAb), except in specific contexts as described below. These terms, as used herein, are meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')<sub>2</sub> fragments) which are capable of specifically binding to a polypeptide or nucleic acid molecule of the invention or a portion thereof. It will therefore be appreciated that, in addition to the intact antibodies of the invention, Fab, F(ab')<sub>2</sub> and other fragments of the antibodies described herein, and other peptides and peptide fragments that bind one or more polypeptides or polynucleotides of the invention, are also encompassed within the scope of the invention. Such antibody fragments are typically produced by proteolytic cleavage of intact antibodies, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Antibody fragments, and peptides or peptide fragments, may also be produced through the application of recombinant DNA technology or through synthetic chemistry.

Epitope-bearing peptides and polypeptides, and nucleic acid molecules or portions thereof, of the invention may be used to induce antibodies according to methods well known in the art, as generally described herein (*see, e.g.*, Sutcliffe, *et al.*, *supra*; Wilson, *et al.*, *supra*; and Bittle, F. J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985)).

Polyclonal antibodies according to this aspect of the invention may be made by immunizing an animal with one or more of the polypeptides or nucleic acid molecules of the invention described herein or portions thereof according to standard techniques (*see, e.g.*, Harlow, E., and Lane, D., *Antibodies: A*

*Laboratory Manual*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1988); Kaufman, P.B., et al., In: *Handbook of Molecular and Cellular Methods in Biology and Medicine*, Boca Raton, Florida: CRC Press, pp. 468-469 (1995)). For producing antibodies that recognize and bind to the polypeptides or nucleic acid molecules of the invention or portions thereof, animals may be immunized with free peptide or free nucleic acid molecules; however, antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as albumin, KLH, or tetanus toxoid (particularly for producing antibodies against the nucleic acid molecules of the invention or portions thereof; see Harlow and Lane, *supra*, at page 154), or to a solid phase carrier such as a latex or glass microbead. For instance, peptides containing cysteine may be coupled to carrier using a linker such as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice may be immunized with either free (if the polypeptide immunogen is larger than about 25 amino acids in length) or carrier-coupled peptides or nucleic acid molecules, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg peptide, polynucleotide, or carrier protein, and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of antibody which can be detected, for example, by ELISA assay using free peptide or nucleic acid molecule adsorbed to a solid surface. In another approach, cells expressing one or more of the polypeptides or polynucleotides of the invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies, according to routine immunological methods. In yet another method, a preparation of one or more of the polypeptides or polynucleotides of the invention is prepared and purified as described herein, to render it substantially free of natural contaminants. Such a preparation may then be introduced into an animal in order to produce polyclonal antisera of greater specific activity. The titer of antibodies in serum from an immunized animal, regardless of the method of immunization used, may be increased by selection of anti-peptide or anti-polynucleotide antibodies, for

instance, by adsorption to the peptide or polynucleotide on a solid support and elution of the selected antibodies according to methods well known in the art.

In an alternative method, the antibodies of the present invention are monoclonal antibodies (or fragments thereof which bind to one or more of the polypeptides of the invention). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler *et al.*, *Nature* 256:495 (1975); Köhler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Köhler *et al.*, *Eur. J. Immunol.* 6:292 (1976); Hammerling *et al.*, In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a polypeptide or polynucleotide of the invention (or a fragment thereof), or with a cell expressing a polypeptide or polynucleotide of the invention (or a fragment thereof). The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP<sub>2</sub>O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastroenterol.* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding one or more of the polypeptides or nucleic acid molecules of the invention, or fragments thereof. Hence, the present invention also provides hybridoma cells and cell lines producing monoclonal antibodies of the invention, particularly that recognize and bind to one or more of the polypeptides or nucleic acid molecules of the invention.

Alternatively, additional antibodies capable of binding to one or more of the polypeptides of the invention, or fragments thereof, may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, antibodies specific for one or more of the polypeptides or polynucleotides of the invention, prepared as described above, are used to immunize an animal, preferably a mouse. The splenocytes of such an

animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to an antibody specific for one or more of the polypeptides or polynucleotides of the invention can be blocked by polypeptides of the invention themselves. Such antibodies comprise anti-idiotypic antibodies to the antibodies recognizing one or more of the polypeptides or polynucleotides of the invention, and can be used to immunize an animal to induce formation of further antibodies specific for one or more of the polypeptides or polynucleotides of the invention.

5

For use, the antibodies of the invention may optionally be detectably labeled by covalent or non-covalent attachment of one or more labels, including but not limited to chromogenic, enzymatic, radioisotopic, isotopic, fluorescent, toxic, chemiluminescent, or nuclear magnetic resonance contrast agents or other labels.

10

15

Examples of suitable enzyme labels include malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

20

25

Examples of suitable radioisotopic labels include  $^3\text{H}$ ,  $^{111}\text{In}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Fe}$ ,  $^{75}\text{Se}$ ,  $^{152}\text{Eu}$ ,  $^{90}\text{Y}$ ,  $^{67}\text{Cu}$ ,  $^{217}\text{Cl}$ ,  $^{211}\text{At}$ ,  $^{212}\text{Pb}$ ,  $^{47}\text{Sc}$ ,  $^{109}\text{Pd}$ , etc.  $^{111}\text{In}$  is a preferred isotope where *in vivo* imaging is used since it avoids the problem of dehalogenation of the  $^{125}\text{I}$  or  $^{131}\text{I}$ -labeled monoclonal antibody by the liver. In addition, this radionuclide has a more favorable gamma emission energy for imaging (Perkins *et al.*, *Eur. J. Nucl. Med.* 10:296-301 (1985); Carasquillo *et al.*, *J. Nucl. Med.* 28:281-287 (1987)). For example,  $^{111}\text{In}$  coupled to monoclonal antibodies with 1-(*P*-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumorous tissues, particularly the liver, and therefore enhances specificity of tumor localization (Esteban *et al.*, *J. Nucl. Med.* 28:861-870 (1987)).

30

Examples of suitable non-radioactive isotopic labels include  $^{157}\text{Gd}$ ,  $^{55}\text{Mn}$ ,  $^{162}\text{Dy}$ ,  $^{52}\text{Tr}$ , and  $^{56}\text{Fe}$ .

Examples of suitable fluorescent labels include an  $^{152}\text{Eu}$  label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a

-91-

phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, a green fluorescent protein (GFP) label, and a fluorescamine label.

Examples of suitable toxin labels include diphtheria toxin, ricin, and cholera toxin.

5 Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

10 Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

15 Typical techniques for binding the above-described labels to the antibodies of the invention are provided by Kennedy *et al.*, *Clin. Chim. Acta* 70:1-31 (1976), and Schurs *et al.*, *Clin. Chim. Acta* 81:1-40 (1977). Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

20 It will be appreciated by one of ordinary skill that the antibodies of the present invention may alternatively be coupled to a solid support, to facilitate, for example, chromatographic and other immunological procedures using such solid phase-immobilized antibodies. Included among such procedures are the use of the antibodies of the invention to isolate or purify polypeptides comprising one or more epitopes encoded by the nucleic acid molecules of the invention (which may be fusion polypeptides or other polypeptides of the invention described herein), or to isolate or purify polynucleotides comprising one or more recombination site sequences of the invention or portions thereof. Methods for isolation and purification of polypeptides (and, by analogy, polynucleotides) by affinity chromatography, for example using the antibodies of the invention coupled to a solid phase support, are well-known in the art and will be familiar to one of ordinary skill. The antibodies of the invention may also be used in other applications, for example to cross-link or couple two or more proteins, polypeptides, polynucleotides, or portions thereof into a structural and/or functional complex. In one such use, an antibody of the invention may have two

25

30

or more distinct epitope-binding regions that may bind, for example, a first polypeptide (which may be a polypeptide of the invention) at one epitope-binding region on the antibody and a second polypeptide (which may be a polypeptide of the invention) at a second epitope-binding region on the antibody, thereby bringing the first and second polypeptides into close proximity to each other such that the first and second polypeptides are able to interact structurally and/or functionally (as, for example, linking an enzyme and its substrate to carry out enzymatic catalysis, or linking an effector molecule and its receptor to carry out or induce a specific binding of the effector molecule to the receptor or a response to the effector molecule mediated by the receptor). Additional applications for the antibodies of the invention include, for example, the preparation of large-scale arrays of the antibodies, polypeptides, or nucleic acid molecules of the invention, or portions thereof, on a solid support, for example to facilitate high-throughput screening of protein or RNA expression by host cells containing nucleic acid molecules of the invention (known in the art as "chip array" protocols; *see, e.g.*, U.S. Patent Nos. 5,856,101, 5,837,832, 5,770,456, 5,744,305, 5,631,734, and 5,593,839, which are directed to production and use of chip arrays of polypeptides (including antibodies) and polynucleotides, and the disclosures of which are incorporated herein by reference in their entireties). By "solid support" is intended any solid support to which an antibody can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polycarbonate, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Preferred are beads made of glass, latex or a magnetic material. Linkage of an antibody of the invention to a solid support can be accomplished by attaching one or both ends of the antibody to the support. Attachment may also be made at one or more internal sites in the antibody. Multiple attachments (both internal and at the ends of the antibody) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulphydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments, addition of an affinity tag sequence to the peptide can be used such as GST

(Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., et al., *J. Chromatog.* 411:77 (1987)), or biotin. Alternatively, attachment can be accomplished using a ligand which binds the Fc region of the antibodies of the invention, e.g., protein A or protein G. Such affinity tags may be used for the reversible attachment of the antibodies to the support. Peptides may also be recognized via specific ligand-receptor interactions or using phage display methodologies that will be familiar to the skilled artisan, for their ability to bind polypeptides of the invention or fragments thereof.

5

10

### Kits

15

20

25

30

In another aspect, the invention provides kits which may be used in producing the nucleic acid molecules, polypeptides, vectors, host cells, and antibodies, and in the recombinational cloning methods, of the invention. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more of the nucleic acid molecules, primers, polypeptides, vectors, host cells, or antibodies of the invention. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins (e.g., Int) or auxiliary factors (e.g. IHF and/or Xis) or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix) one or more Destination Vector molecules (including those described herein), one or more Entry Clone or Entry Vector molecules (including those described herein), one or more primer nucleic acid molecules (particularly those described herein), one or more host cells (e.g. competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, and the corresponding U.S. Utility Application No. \_\_\_\_\_ of Hartley et al., entitled "Cells Resistant to Toxic Genes and Uses Thereof," filed

on even day herewith, the disclosures of which are incorporated by reference herein in its entirety), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, such as one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites (or portions thereof) of the invention, and particularly one or more of the nucleic acid molecules contained in the deposited clones described herein. Kits according to this aspect of the invention may also comprise one or more isolated nucleic acid molecules of the invention, one or more vectors of the invention, one or more primer nucleic acid molecules of the invention, and/or one or more antibodies of the invention. The kits of the invention may further comprise one or more additional containers containing one or more additional components useful in combination with the nucleic acid molecules, polypeptides, vectors, host cells, or antibodies of the invention, such as one or more buffers, one or more detergents, one or more polypeptides having nucleic acid polymerase activity, one or more polypeptides having reverse transcriptase activity, one or more transfection reagents, one or more nucleotides, and the like. Such kits may be used in any process advantageously using the nucleic acid molecules, primers, vectors, host cells, polypeptides, antibodies and other compositions of the invention, for example in methods of synthesizing nucleic acid molecules (e.g., via amplification such as via PCR), in methods of cloning nucleic acid molecules (preferably via recombinational cloning as described herein), and the like.

#### *Optimization of Recombinational Cloning System*

The usefulness of a particular nucleic acid molecule, or vector comprising a nucleic acid molecule, of the invention in methods of recombinational cloning may be determined by any one of a number of assay methods. For example, Entry and Destination vectors of the present invention may be assessed for their ability to function (*i.e.*, to mediate the transfer of a nucleic acid molecule, DNA segment, gene, cDNA molecule or library from a cloning vector to an Expression Vector) by carrying out a recombinational cloning reaction as described in more detail in the Examples below and as described in U.S. Application Nos. 08/663,002, filed

5

10

June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of which are incorporated by reference herein in their entireties. Alternatively, the functionality of Entry and Destination Vectors prepared according to the invention may be assessed by examining the ability of these vectors to recombine and create cointegrate molecules, or to transfer a nucleic acid molecule of interest, using an assay such as that described in detail below in Example 19. Analogously, the formulation of compositions comprising one or more recombination proteins or combinations thereof, for example GATEWAY™ LR Clonase™ Enzyme Mix and GATEWAY™ BP Clonase™ Enzyme Mix, may be optimized using assays such as those described below in Example 18.

20

25

30

### Uses

There are a number of applications for the compositions, methods and kits of the present invention. These uses include, but are not limited to, changing vectors, targeting gene products to intracellular locations, cleaving fusion tags from desired proteins, operably linking nucleic acid molecules of interest to regulatory genetic sequences (*e.g.*, promoters, enhancers, and the like), constructing genes for fusion proteins, changing copy number, changing replicons, cloning into phages, and cloning, *e.g.*, PCR products, genomic DNAs, and cDNAs. In addition, the nucleic acid molecules, vectors, and host cells of the invention may be used in the production of polypeptides encoded by the nucleic acid molecules, in the production of antibodies directed against such polypeptides, in recombinational cloning of desired nucleic acid sequences, and in other applications that may be enhanced or facilitated by the use of the nucleic acid molecules, vectors, and host cells of the invention.

In particular, the nucleic acid molecules, vectors, host cells, polypeptides, antibodies, and kits of the invention may be used in methods of transferring one or more desired nucleic acid molecules or DNA segments, for example one or more genes, cDNA molecules or cDNA libraries, into a cloning or Expression Vector for use in transforming additional host cells for use in cloning or

amplification of, or expression of the polypeptide encoded by, the desired nucleic acid molecule or DNA segment. Such recombinational cloning methods which may advantageously use the nucleic acid molecules, vectors, and host cells of the invention, are described in detail in the Examples below, and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of all of which are incorporated by reference herein in their entireties.

10

It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent from the description of the invention contained herein in view of information known to the ordinarily skilled artisan, and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

15

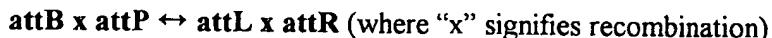
### *Examples*

#### *Example 1: Recombination Reactions of Bacteriophage λ*

The *E. coli* bacteriophage λ can grow as a lytic phage, in which case the host cell is lysed, with the release of progeny virus. Alternatively, lambda can integrate into the genome of its host by a process called lysogenization (see Figure 60). In this lysogenic state, the phage genome can be transmitted to daughter cells for many generations, until conditions arise that trigger its excision from the genome. At this point, the virus enters the lytic part of its life cycle. The control of the switch between the lytic and lysogenic pathways is one of the best understood processes in molecular biology (M. Ptashne, *A Genetic Switch*, Cell Press, 1992).

The integrative and excisive recombination reactions of  $\lambda$ , performed *in vitro*, are the basis of Recombinational Cloning System of the present invention. They can be represented schematically as follows:

5



10

The four att sites contain binding sites for the proteins that mediate the reactions. The wild type attP, attB, attL, and attR sites contain about 243, 25, 100, and 168 base pairs, respectively. The attB x attP reaction (hereinafter referred to as a "BP Reaction," or alternatively and equivalently as an "Entry Reaction" or a "Gateward Reaction") is mediated by the proteins Int and IHF. The attL x attR reaction (hereinafter referred to as an "LR Reaction," or alternatively and equivalently as a "Destination Reaction") is mediated by the proteins Int, IHF, and Xis. Int (integrase) and Xis (excisionase) are encoded by the  $\lambda$  genome, while IHF (integration host factor) is an *E. coli* protein. For a general review of lambda recombination, see: A. Landy, *Ann. Rev. Biochem.* 58: 913-949 (1989).

15

20

**Example 2: Recombination Reactions of the Recombinational Cloning System**

The LR Reaction -- the exchange of a DNA segment from an Entry Clone to a Destination Vector -- is the *in vitro* version of the  $\lambda$  excision reaction:

25



30

There is a practical imperative for this configuration: after an LR Reaction in one configuration of the present method, an att site usually separates a functional motif (such as a promoter or a fusion tag) from a nucleic acid molecule of interest in an Expression Clone, and the 25 bp attB site is much smaller than the attP, attL, and attR sites.

Note that the recombination reaction is conservative, i.e., there is no net synthesis or loss of base pairs. The DNA segments that flank the recombination

sites are merely switched. The wild type  $\lambda$  recombination sites are modified for purposes of the GATEWAY™ Cloning System, as follows:

To create certain preferred Destination Vectors, a part (43 bp) of attR was removed, to make the excisive reaction irreversible and more efficient (W. Bushman et al., *Science* 230: 906, 1985). The attR sites in preferred Destination Vectors of the invention are 125 bp in length. Mutations were made to the core regions of the att sites, for two reasons: (1) to eliminate stop codons, and (2) to ensure specificity of the recombination reactions (i.e., attR1 reacts only with attL1, attR2 reacts only with attL2, etc.).

Other mutations were introduced into the short (5 bp) regions flanking the 15 bp core regions of the attB sites to minimize secondary structure formation in single-stranded forms of attB plasmids, e.g., in phagemid ssDNA or in mRNA. Sequences of attB1 and attB2 to the left and right of a nucleic acid molecule of interest after it has been cloned into a Destination Vector are given in Figure 6.

Figure 61 illustrates how an Entry Clone and a Destination Vector recombine in the LR Reaction to form a co-integrate, which resolves through a second reaction into two daughter molecules. The two daughter molecules have the same general structure regardless of which pair of sites, attL1 and attR1 or attL2 and attR2, react first to form the co-integrate. The segments change partners by these reactions, regardless of whether the parental molecules are both circular, one is circular and one is linear, or both are linear. In this example, selection for ampicillin resistance carried on the Destination Vector, which also carries the death gene ccdB, provides the means for selecting only for the desired attB product plasmid.

### ***Example 3: Protein Expression in the Recombinational Cloning System***

Proteins are expressed *in vivo* as a result of two processes, transcription (DNA into RNA), and translation (RNA into protein). For a review of protein expression in prokaryotes and eukaryotes, see Example 13 below. Many vectors (pUC, BlueScript, pGem) use interruption of a transcribed lacZ gene for blue-white screening. These plasmids, and many Expression Vectors, use the lac promoter to control expression of cloned genes. Transcription from the lac

promoter is turned on by adding the inducer IPTG. However, a low level of RNA is made in the absence of inducer, i.e., the lac promoter is never completely off. The result of this "leakiness" is that genes whose expression is harmful to *E. coli* may prove difficult or impossible to clone in vectors that contain the lac promoter, or they may be cloned only as inactive mutants.

In contrast to other gene expression systems, nucleic acid molecules cloned into an Entry Vector may be designed *not* to be expressed. The presence of the strong transcriptional terminator *rrnB* (Orosz, et al., *Eur. J. Biochem.* 201: 653, 1991) just upstream of the attL1 site keeps transcription from the vector promoters (drug resistance and replication origin) from reaching the cloned gene. However, if a toxic gene is cloned into a Destination Vector, the host may be sick, just as in other expression systems. But the reliability of subcloning by *in vitro* recombination makes it easier to recognize that this has happened -- and easier to try another expression option in accordance with the methods of the invention, if necessary.

#### ***Example 4: Choosing the Right Entry Vector***

There are two kinds of choices that must be made in choosing the best Entry Vector, dictated by (1) the particular DNA segment that is to be cloned, and (2) what is to be accomplished with the cloned DNA segment. These factors are critical in the choice of Entry Vector used, because when the desired nucleic acid molecule of interest is moved from the Entry Vector to a Destination Vector, all the base pairs between the nucleic acid molecule of interest and the Int cutting sites in attL1 and attL2 (such as in Figure 6) move into the Destination Vector as well. For genomic DNAs that are not expressed as a result of moving into a Destination Vector, these decisions are not as critical.

For example, if an Entry Vector with certain translation start signals is used, those sequences will be translated into amino acids if an amino-terminal fusion to the desired nucleic acid molecule of interest is made. Whether the desired nucleic acid molecule of interest is to be expressed as fusion protein, native protein, or both, dictates whether translational start sequences must be included between the attB sites of the clone (native protein) or, alternatively, supplied by the Destination

-100-

Vector (fusion protein). In particular, Entry Clones that include translational start sequences may prove less suitable for making fusion proteins, as internal initiation of translation at these sites can decrease the yield of N-terminal fusion protein. These two types of expression afforded by the compositions and methods of the invention are illustrated in Figure 62.

No Entry Vector is likely to be optimal for all applications. The nucleic acid molecule of interest may be cloned into any of several optimal Entry Vectors.

As an example, consider pENTR7 (Figure 16) and pENTR11 (Figure 20), which are useful in a variety of applications, including (but not limited to):

10

- Cloning cDNAs from most of the commercially available libraries. The sites to the left and right of the *ccdB* death gene have been chosen so that directional cloning is possible if the DNA to be cloned does not have two or more of these restriction sites.

15

- Cloning of genes directionally: *SaII*, *BamHI*, *XmnI* (blunt), or *KpnI* on the left of *ccdB*; *NotI*, *XhoI*, *XbaI*, or *EcoRV* (blunt), on the right.

20

- Cloning of genes or gene fragments with a blunt amino end at the *XmnI* site. The *XmnI* site has four of the six most favored bases for eukaryotic expression (see Example 13, below), so that if the first three bases of the DNA to be cloned are ATG, the open reading frame (ORF) will be expressed in eukaryotic cells (e.g., mammalian cells, insect cells, yeast cells) when it is transcribed in the appropriate Destination Vector. In addition, in pENTR11, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

25

- Cleaving off amino terminal fusions (e.g., His<sub>6</sub>, GST, or thioredoxin) using the highly specific TEV (Tobacco Etch Virus) protease (available from Life Technologies, Inc.). If the nucleic acid molecule of interest is cloned at the

30

-101-

blunt *XmnI* site, TEV cleavage will leave two amino acids on the amino end of the expressed protein.

5           • Selecting against uncut or singly cut Entry Vector molecules during cloning with restriction enzymes and ligase. If the *ccdB* gene is not removed with a double digest, it will kill any recipient *E. coli* cell that does not contain a mutation that makes the cell resistant to *ccdB* (see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety).

10

• Allowing production of amino fusions with ORFs in all cloning sites. There are no stop codons (in the attL1 reading frame) upstream of the *ccdB* gene.

In addition, pENTR11 is also useful in the following applications:

15

• Cloning cDNAs that have an *NcoI* site at the initiating ATG into the *NcoI* site. Similar to the *XmnI* site, this site has four of the six most favored bases for eukaryotic expression. Also, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

20

• Producing carboxy fusion proteins with ORFs positioned in phase with the reading frame convention for carboxy-terminal fusions (see Figure 20A).

25           Table 1 lists some non-limiting examples of Entry Vectors and their characteristics, and Figures 10-20 show their cloning sites. All of the Entry Vectors listed in Table 1 are available commercially from Life Technologies, Inc., Rockville, Maryland. Other Entry Vectors not specifically listed here, which comprise alternative or additional features may be made by one of ordinary skill using routine methods of molecular and cellular biology, in view of the disclosure contained herein.

30

Table 1 Examples of Entry Vectors

Designation	Mnemonic Name	Class of Entry Vector	Distinctive Cloning Sites	Amino Fusions	Native Protein in E.coli	Native Protein in Eukaryotic Cells	Protein Synthesis Features
pENTR-1A, 2B, 3C	Minimal blunt RF A, B, C	Alternative Reading Frame Vectors	Reading frame A, B, or C; blunt cut closest to attL1	Good	Poor	Good	Minimal amino acids between tag and protein; no SD
pENTR4	Minimal Nco	Restr. Enz. Cleavage Vectors	Nco I site (common in euk. cDNAs) closest to attL1	Good	Poor	Good	Good Kozac; no SD
pENTR5	Minimal Nde	Restr. Enz. Cleavage Vectors	Nde I site closest to attL1	Good	Poor	Poor at Nde I, Good at Xmn I	No SD; poor Kozac at Nde, good at Xmn
pENTR6	Minimal Sph	Restr. Enz. Cleavage Vectors	Sph I site closest to attL1	Good	Poor	Poor at Sph I, Good at Xmn I	No SD; poor Kozac at Sph, good at Xmn
pENTR7	TEV Blunt	TEV Cleavage Site Present	Xmn I (blunt) is first cloning site after TEV site	Good	Poor	Good at Xmn I site	TEV protease leaves Gly-Thr on amino end of protein; no SD
pENTR8	TEV Nco	TEV Cleavage Site Present	Nco I is first cloning site after TEV site	Good	Poor	Good	TEV protease leaves Gly-Thr on amino end of protein; no SD

-103-

pENTR9	TEV Nde	TEV Cleavage Site Present	Nde I is first cloning site after TEV site	Good	Poor	Poor	TEV protease leaves Gly-Thr on amino end of protein; no SD, poor Kozac
pENTR10	Nde with SD	Good SD for E.coli Expression	Strong SD; Nde I site, no TEV	Poor	Good	Poor	Strong SD, internal starts in amino fusions. Poor Kz. No TEV
pENTR11	2 X SD+Kozac	Good SD for E.coli Expression	Xmn I (blunt) and Nco I sites each preceded by SD and Kozac	Good	Good	Good	Strong SD/Koz Internal starts in amino fusions. No TEV

Entry vectors pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), and pENTR3C (Figures 12A and 12B) are almost identical, except that the restriction sites are in different reading frames. Entry vectors pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), and pENTR6 (Figures 15A and 15B) are essentially identical to pENTR1A, except that the blunt *Dra*I site has been replaced with sites containing the ATG methionine codon: *Nco*I in pENTR4, *Nde*I in pENTR5, and *Sph*I in pENTR6. Nucleic acid molecules that contain one of these sites at the initiating ATG can be conveniently cloned in these Entry vectors. The *Nco*I site in pENTR4 is especially useful for expression of nucleic acid molecules in eukaryotic cells, since it contains many of the bases that give efficient translation (see Example 13, below). (Nucleic acid molecules of interest cloned into the *Nde*I site of pENTR5 are not expected to be highly expressed in eukaryotic cells, because the cytosine at position -3 from the initiating ATG is rare in eukaryotic genes.)

Entry vectors pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), and pENTR9 (Figures 18A and 18B) contain the recognition site for the TEV protease between the attL1 site and the cloning sites. Cleavage sites for *Xmn*I (blunt), *Nco*I, and *Nde*I, respectively, are the most 5' sites in these Entry vectors. Amino fusions can be removed efficiently if nucleic acid molecules are cloned into these Entry vectors. TEV protease is highly active and highly specific.

#### *Example 5: Controlling Reading Frame*

One of the trickiest tasks in expression of cloned nucleic acid molecules is making sure the reading frame is correct. (Reading frame is important if fusions are being made between two ORFs, for example between a nucleic acid molecule of interest and a His6 or GST domain.) For purposes of the present invention, the following convention has been adopted: The reading frame of the DNA cloned into any Entry Vector must be in phase with that of the attB1 site shown in Figure 16A, pENTR7. Notice that the six As of the attL1 site are split into two lysine codons (aaa aaa). The Destination Vectors that make amino fusions were constructed such that they enter the attR1 site in this reading frame.

-105-

Destination Vectors for carboxy terminal fusions were also constructed, including those containing His<sub>6</sub> (pDEST23; Figure 43), GST (pDEST24; Figure 44), or thioredoxin (pDEST25; Figure 45) C-terminal fusion sequences.

Therefore, if a nucleic acid molecule of interest is cloned into an Entry Vector so that the aaa aaa reading frame within the attL1 site is in phase with the nucleic acid molecule's ORF, amino terminal fusions will automatically be correctly phased, for all the fusion tags. This is a significant improvement over the usual case, where each different vector can have different restriction sites and different reading frames.

See Example 15 for a practical example of how to choose the most appropriate combinations of Entry Vector and Destination Vector.

## Materials

Unless otherwise indicated, the following materials were used in the remaining Examples included herein:

### 5X LR Reaction Buffer:

200-250 mM (preferably 250 mM) Tris-HCl, pH 7.5  
250-350 mM (preferably 320 mM) NaCl  
1.25-5 mM (preferably 4.75 mM) EDTA  
12.5-35 mM (preferably 22-35 mM, and most preferably 35 mM)  
Spermidine-HCl  
1 mg/ml bovine serum albumin

### GATEWAY™ LR Clonase™ Enzyme Mix:

per 4 µl of 1X LR Reaction Buffer:

150 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

-106-

25 ng carboxy-His6-tagged Xis (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

30 ng IHF

5 50% glycerol

**5X BP Reaction Buffer:**

125 mM Tris-HCl, pH 7.5

110 mM NaCl

10 25 mM EDTA

25 mM Spermidine-HCl

5 mg/ml bovine serum albumin

**GATEWAY™ BP Clonase™ Enzyme Mix:**

15 per 4 µl of 1X BP Reaction Buffer:

200 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

80 ng IHF

20 50% glycerol

**10X Clonase Stop Solution:**

25 50 mM Tris-HCl, pH 8.0

1 mM EDTA

2 mg/ml Proteinase K

***Example 6: LR ("Destination") Reaction***

To create a new Expression Clone containing the nucleic acid molecule of interest (and which may be introduced into a host cell, ultimately for production of the polypeptide encoded by the nucleic acid molecule), an Entry Clone or Vector containing the nucleic acid molecule of interest, prepared as described

herein, is reacted with a Destination Vector. In the present example, a  $\beta$ -Gal gene flanked by attL sites is transferred from an Entry Clone to a Destination Vector.

Materials needed:

- 5      • 5 X LR Reaction buffer
- 10     • Destination Vector (preferably linearized), 75-150 ng/ $\mu$ l
- Entry Clone containing nucleic acid molecule of interest, 100-300 ng in  $\leq$  8  $\mu$ l  
          TE buffer
- Positive control Entry Clone (pENTR- $\beta$ -Gal) DNA (See note, below)
- 15     • Positive control Destination Vector, pDEST1 (pTrc), 75 ng/ $\mu$ l
- GATEWAY™ LR Clonase™ Enzyme Mix (stored at - 80° C)
- 10X Clonase Stop solution
- pUC19 DNA, 10 pg/ $\mu$ l
- Chemically competent *E. coli* cells (competence:  $\geq 1 \times 10^7$  CFU/ $\mu$ g), 400  $\mu$ l.
- 15     • LB Plates containing ampicillin (100  $\mu$ g/ml) and methicillin (200  $\mu$ g/ml)  $\pm$   
          X-gal and IPTG (See below)

Notes:

Preparation of the Entry Clone DNA: Miniprep DNA that has been treated with RNase works well. A reasonably accurate quantitation ( $\pm 50\%$ ) of the DNA to be cloned is advised, as the GATEWAY™ reaction appears to have an optimum of about 100-300 ng of Entry Clone per 20  $\mu$ l of reaction mix.

The positive control Entry Clone, pENTR- $\beta$ -Gal, permits functional analysis of clones based on the numbers of expected blue vs. white colonies on LB plates containing IPTG + Bluo-gal (or X-gal), in addition to ampicillin (100  $\mu$ g/ml) and methicillin (200  $\mu$ g/ml). Because  $\beta$ -Galactosidase is a large protein, it often yields a less prominent band than many smaller proteins do on SDS protein gels.

In the Positive Control Entry Vector pENTR- $\beta$ -Gal, the coding sequence of  $\beta$ -Gal has been cloned into pENTR11 (Figures 20A and 20B), with translational start signals permitting expression in *E. coli*, as well as in eukaryotic

-108-

cells. The positive control Destination Vector, for example pDEST1 (Figure 21), is preferably linearized.

To prepare X-gal + IPTG plates, either of the following protocols may be used:

5

A. With a glass rod, spread over the surface of an LB agar plate: 40 µl of 20 mg/ml X-gal (or Bluo-gal) in DMF plus 4 µl 200 mg/ml IPTG. Allow liquid to adsorb into agar for 3-4 hours at 37° C before plating cells.

10

B. To liquid LB agar at ~45°C, add: X-gal (or Bluo-Gal) (20 mg/ml in DMF) to make 50 µg/ml and IPTG (200 mM in water) to make 0.5-1 mM, just prior to pouring plates. Store X-gal and Bluo-Gal in a light-shielded container.

15

Colony color may be enhanced by placing the plates at 5°C for a few hours after the overnight incubation at 37°C. Protocol B can give more consistent colony color than A, but A is more convenient when selection plates are needed on short notice.

20

Recombination in Clonase reactions continues for many hours. While incubations of 45-60 minutes are usually sufficient, reactions with large DNAs, or in which both parental DNAs are supercoiled, or which will be transformed into cells of low competence, can be improved with longer incubation times, such as 2-24 hours at 25°C.

Procedure:

25

1. Assemble reactions as follows (combine all components at room temperature, except GATEWAY™ LR Clonase™ Enzyme Mix ("Clonase LR"), before removing Clonase LR from frozen storage):

		Tube 1	Tube 2	Tube 3	Tube 4
	Component	Neg.	Pos.	Neg.	Test
5	p-Gate-βGal, (Positive control Entry Clone) 75 ng/μl	4 μl	4 μl		
10	pDEST1 (Positive control Destination Vector), 75 ng/μl	4 μl	4 μl		
15	Your Entry Clone (100-300 ng)			1 - 8 μl	1 - 8 μl
20	Destination Vector for your nucleic acid molecule, 75 ng/μl			4 μl	4 μl
	5 X LR Reaction Buffer	4 μl	4 μl	4 μl	4 μl
	TE	8 μl	4 μl	To 20 μl	To 16 μl
	GATEWAY™ LR Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μl	---	4 μl
	Total Volume	20 μl	20 μl	20 μl	20 μl

2. Remove the GATEWAY™ LR Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.

3. Add 4 μl of GATEWAY™ LR Clonase™ Enzyme Mix to reactions #2 and #4;

25 4. Return GATEWAY™ LR Clonase™ Enzyme Mix to - 80° C freezer.

5. Incubate tubes at 25° for at least 60 minutes.

6. Add 2 μl Clonase Stop solution to all reactions. Incubate for 20 min at 37°C. (This step usually increases the total number of colonies obtained by 10-20 fold.)

7. Transform 2 μl into 100 μl competent *E. coli*. Select on plates containing ampicillin at 100 μg/ml.

**Example 7: Transformation of *E. coli***

To introduce cloning or Expression Vectors prepared using the recombinational cloning system of the invention, any standard *E. coli* transformation protocol should be satisfactory. The following steps are recommended for best results:

-110-

1. Let the mixture of competent cells and Recombinational Cloning System reaction product stand on ice at least 15 minutes prior to the heat-shock step. This gives time for the recombination proteins to dissociate from the DNA, and improves the transformation efficiency.

5

2. Expect the reaction to be about 1%-5% efficient, i.e., 2 µl of the reaction should contain at least 100 pg of the Expression Clone plasmid (taking into account the amounts of each parental plasmid in the reaction, and the subsequent dilution). If the E. coli cells have a competence of  $10^7$  CFU/µg, 100 pg of the desired clone plasmid will give about 1000 colonies, or more, if the entire transformation is spread on one ampicillin plate.

10

3. Always do a control pUC DNA transformation. If the number of colonies is not what you expect, the pUC DNA transformation gives you an indication 15 of where the problem was.

15

***Example 8: Preparation of attB-PCR Product***

For preparation of attB-PCR products in the PCR cloning methods described in Example 9 below, PCR primers containing attB1 and attB2 sequences are used. The attB1 and attB2 primer sequences are as follows:

**attB1:** 5'-GGGGACAAGTTGTACAAAAAAGCAGGCT-(template-specific sequence)-3'

**attB2:** 5'-GGGGACCCTTGTACAAGAAAGCTGGGT-(template-specific sequence)-3'

The attB1 sequence should be added to the amino primer, and the attB2 sequence to the carboxy primer. The 4 guanines at the 5' ends of each of these primers 30 enhance the efficiency of the minimal 25 bp attB sequences as substrates for use in the cloning methods of the invention.

Standard PCR conditions may be used to prepare the PCR product. The following suggested protocol employs PLATINUM *Taq* DNA Polymerase High

-111-

Fidelity®, available commercially from Life Technologies, Inc. (Rockville, MD). This enzyme mix eliminates the need for hot starts, has improved fidelity over Taq, and permits synthesis of a wide range of amplicon sizes, from 200 bp to 10 kb, or more, even on genomic templates.

5

### **Materials needed:**

- PLATINUM Taq DNA Polymerase High Fidelity® (Life Technologies, Inc.)
- attB1- and attB2- containing primer pair (see above) specific for your template
- DNA template (linearized plasmid or genomic DNA)
- 10X High Fidelity PCR Buffer
- 10 mM dNTP mix
- PEG/MgCl<sub>2</sub> Mix (30% PEG 8000, 30 mM MgCl<sub>2</sub>)

10

15

### Procedure:

1.) Assemble the reaction as follows:

Component	Reaction with <u>Plasmid Target</u>	Reaction with <u>Genomic Target</u>
10X High Fidelity PCR Buffer	5 µl	5 µl
dNTP Mix 10 mM	1 µl	1 µl
MgSO <sub>4</sub> , 50mM	2 µl	2 µl
attB1 Primer, 10 µM	2 µl	1 µl
attB2 Primer, 10 µM	2 µl	1 µl
Template DNA	1-5 ng*	≥100 ng
PLATINUM Taq High Fidelity	2 µl	1 µl
Water	to 50 µl	to 50 µl

\* Use of higher amounts of plasmid template may permit fewer cycles (10-15) of PCR

30

35

2.) Add 2 drops mineral oil, as appropriate.

3.) Denature for 30 sec. at 94°C.

4.) Perform 25 cycles:

5           94°C for 15 sec-30 sec

55°C for 15 sec-30 sec

68°C for 1 min per kb of template.

10           5.) Following the PCR reaction, apply 1-2 µl of the reaction mixture to an agarose gel, together with size standards (e.g., 1 Kb Plus Ladder, Life Technologies, Inc.) and quantitation standards (e.g., Low Mass Ladder, Life Technologies, Inc.), to assess the yield and uniformity of the product.

15           Purification of the PCR product is recommended, to remove attB primer dimers which can clone efficiently into the Entry Vector. The following protocol is fast and will remove DNA <300 bp in size:

16           6.) Dilute the 50 µl PCR reaction to 200 µl with TE.

17           7.) Add 100 µl PEG/MgCl<sub>2</sub> Solution. Mix and centrifuge immediately at 13,000 RPM for 10 min at room temperature. Remove the supernatant (pellet is clear and hard to see).

18           8.) Dissolve the pellet in 50 µl TE and check recovery on a gel.

19           If the starting PCR template is a plasmid that contains the gene for Kan<sup>r</sup>, it is advisable to treat the completed PCR reaction with the restriction enzyme *Dpn*I, to degrade the plasmid since unreacted residual starting plasmid is a potential source of false-positive colonies from the transformation of the GATEWAY™ Cloning System reaction. Adding ~5 units of *Dpn*I to the completed PCR reaction and incubating for 15 min at 37°C will eliminate this potential problem. Heat inactivate the *Dpn*I at 65°C for 15 min, prior to using the PCR product in the GATEWAY™ Cloning System reaction.

**Example 9: Cloning attB-PCR products into Entry Vectors via the BP ("Gateway") Reaction**

The addition of 5'-terminal attB sequences to PCR primers allows synthesis of a PCR product that is an efficient substrate for recombination with a Donor (attP) Plasmid in the presence of GATEWAY™ BP Clonase™ Enzyme Mix. This reaction produces an Entry Clone of the PCR product (See Figure 8).

The conditions of the Gateway Cloning reaction with an attB PCR substrate are similar to those of the BP Reaction (see Example 10 below), except that the attB-PCR product (see Example 8) substitutes for the Expression Clone, and the attB-tet' PCR product positive control (attB-tet') substitutes for the Expression Clone Positive Control (GFP).

**Materials needed:**

- 5 X BP Reaction Buffer
- Desired attB-PCR product DNA, 50-100 ng in ≤ 8 µl TE.
- Donor (attP) Plasmid (Figures 49-54), 75 ng/µl, supercoiled DNA
- attB-tet' PCR product positive control, 25 ng/µl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at - 80° C)
- 10x Clonase Stop Solution
- pUC19 DNA, 10 pg/µl.
- Chemically competent E.coli cells (competence: ≥ 1×10<sup>7</sup> CFU/µg), 400 µl

**Notes:**

- Preparation of attB-PCR DNA: see Example 8.

- The Positive Control attB-tet' PCR product contains a functional copy of the tet' gene of pBR322, with its own promoter. By plating the transformation of the control BP Reaction on kanamycin (50 µg/ml) plates (if kan' Donor Plasmids are used; see Figures 49-52) or an alternative selection agent (e.g., gentamycin, if gen' Donor Plasmids are used; see Figure 54), and then picking about 50 of these colonies onto plates with tetracycline (20 µg/ml), the

percentage of Entry Clones containing functional tet<sup>r</sup> among the colonies from the positive control reaction can be determined (% Expression Clones = (number of tet<sup>r</sup> + kan<sup>r</sup> (or gen<sup>r</sup>) colonies/kan<sup>r</sup> (or gen<sup>r</sup>) colonies).

5

**Procedure:**

1. Assemble reactions as follows. Combine all components except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from frozen storage.

10

Component	Neg.	Pos.	Test
	Tube 1	Tube 2	Tube 3
attB-PCR product, 50-100 ng			1 - 8 µl
Donor (attP) Plasmid 75 ng/µl	2 µl	2 µl	2 µl
attB-PCR tet <sup>r</sup> control DNA (75 ng/µl)		4 µl	
5 X BP Reaction Buffer	4 µl	4 µl	4 µl
TE	10 µl	6 µl	To 16 µl
GATEWAY™ BP Clonase™ Enzyme Mix (store at -80° C, add last)	4 µl	4 µl	4 µl
Total Volume	20 µl	20 µl	20 µl

15  
20  
25  
30

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.  
 3. Add 4 µl of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.  
 4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.  
 5. Incubate tubes at 25° for at least 60 minutes.

6. Add 2 µl Proteinase K (2 µg/µl) to all reactions. Incubate for 20 min at 37°C.
7. Transform 2 µl into 100 µl competent E. coli, as per 3.2, above. Select on LB plates containing kanamycin, 50 µg/ml.

5

Results:

In initial experiments, primers for amplifying tetR and ampR from pBR322 were constructed containing only the tetR- or ampR-specific targeting sequences, the targeting sequences plus attB1 (for forward primers) or attB2 (for reverse primers) sequences shown in Figure 9, or the attB1 or attB2 sequences with a 5' tail of four guanines. The construction of these primers is depicted in Figure 65. After PCR amplification of tetR and ampR from pBR322 using these primers and cloning the PCR products into host cells using the recombinational cloning system of the invention, the results shown in Figure 66 were obtained. These results demonstrated that primers containing attB sequences provided for a somewhat higher number of colonies on the tetracycline and ampicillin plates. However, inclusion of the 5' extensions of four or five guanines on the primers in addition to the attB sequences provided significantly better cloning results, as shown in Figures 66 and 67. These results indicate that the optimal primers for cloning of PCR products using recombinational cloning will contain the recombination site sequences with a 5' extension of four or five guanine bases.

10

15

20

25

30

To determine the optimal stoichiometry between attB-containing PCR products and attP-containing Donor plasmid, experiments were conducted where the amount of PCR product and Donor plasmid were varied during the BP Reaction. Reaction mixtures were then transformed into host cells and plated on tetracycline plates as above. Results are shown in Figure 68. These results indicate that, for optimal recombinational cloning results with a PCR product in the size range of the tet gene, the amounts of attP-containing Donor plasmids are between about 100-500 ng (most preferably about 200-300 ng), while the optimal concentrations of attB-containing PCR products is about 25-100 ng (most preferably about 100 ng), per 20 µl reaction.

Experiments were then conducted to examine the effect of PCR product size on efficiency of cloning via the recombinational cloning approach of the invention.

PCR products containing attB1 and attB2 sites, at sizes 256 bp, 1 kb, 1.4 kb, 3.4 kb, 4.6 kb, 6.9 kb and 10.1 kb were prepared and cloned into Entry vectors as described above, and host cells were transformed with the Entry vectors containing the cloned PCR products. For each PCR product, cloning efficiency was calculated relative to cloning of pUC19 positive control plasmids as follows:

$$\text{Cloning Efficiency} = \frac{\text{CFU/ng attB PCR product}}{\text{CFU/ng pUC19 control}} \times \frac{\text{Size (kb) PCR product}}{\text{Size (kb) pUC19 control}}$$

The results of these experiments are depicted in Figures 69A-69C (for 256 bp PCR fragments), 70A-70C (for 1 kb PCR fragments), 71A-71C (for 1.4 kb PCR fragments), 72A-72C (for 3.4 kb PCR fragments), 73A-73C (for 4.6 kb PCR fragments), 74 (for 6.9 kb PCR fragments), and 75-76 (for 10.1 kb PCR fragments). The results shown in these figures are summarized in Figure 77, for different weights and moles of input PCR DNA.

Together, these results demonstrate that attB-containing PCR products ranging in size from about 0.25 kb to about 5 kb clone relatively efficiently in the recombinational cloning system of the invention. While PCR products larger than about 5 kb clone less efficiently (apparently due to slow resolution of cointegrates), longer incubation times during the recombination reaction appears to improve the efficiency of cloning of these larger PCR fragments. Alternatively, it may also be possible to improve efficiency of cloning of large (> about 5 kb) PCR fragments by using lower levels of input attP Donor plasmid and perhaps attB-containing PCR product, and/or by adjusting reaction conditions (e.g., buffer conditions) to favor more rapid resolution of the cointegrates.

#### *Example 10: The BP Reaction*

One purpose of the Gateward ("Entry") reaction is to convert an Expression Clone into an Entry Clone. This is useful when you have isolated an individual Expression Clone from an Expression Clone cDNA library, and you wish to transfer the nucleic acid molecule of interest into another Expression Vector, or

to move a population of molecules from an attB or attL library. Alternatively, you may have mutated an Expression Clone and now wish to transfer the mutated nucleic acid molecule of interest into one or more new Expression Vectors. In both cases, it is necessary first to convert the nucleic acid molecule of interest to an Entry Clone.

5

Materials needed:

- 5 X BP Reaction Buffer
- Expression Clone DNA, 100-300 ng in ≤ 8 µl TE.
- Donor (attP) Vector, 75 ng/µl, supercoiled DNA
- Positive control attB-tet-PCR DNA, 25 ng/µl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at - 80°C)
- Clonase Stop Solution (Proteinase K, 2 µg/µl).

10

15

Notes:

Preparation of the Expression Clone DNA: Miniprep DNA treated with RNase works well.

1. As with the LR Reaction (see Example 14), the BP Reaction is strongly influenced by the topology of the reacting DNAs. In general, the reaction is most efficient when one of the DNAs is linear and the other is supercoiled, compared to reactions where the DNAs are both linear or both supercoiled. Further, linearizing the attB Expression Clone (anywhere within the vector) will usually give more colonies than linearizing the Donor (attP) Plasmid. If finding a suitable cleavage site within your Expression Clone vector proves difficult, you may linearize the Donor (attP) Plasmid between the attP1 and attP2 sites (for example, at the NcoI site), avoiding the ccdB gene. Maps of Donor (attP) Plasmids are given in Figures 49-54.

20

25

Procedure:

30 1. Assemble reactions as follows. Combine all components at room temperature, except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from freezer.

Component	Neg.	Pos.	Test
	Tube 1	Tube 2	Tube 3
Positive Control, attB-tet-PCR DNA, 25 ng/ $\mu$ l	4 $\mu$ l	4 $\mu$ l	
Desired attB Expression Clone DNA (100ng) linearized			1 - 8 $\mu$ l
Donor (attP) Plasmid, 75 ng/ $\mu$ l	2 $\mu$ l	2 $\mu$ l	2 $\mu$ l
5 X BP Reaction Buffer	4 $\mu$ l	4 $\mu$ l	4 $\mu$ l
TE	10 $\mu$ l	6 $\mu$ l	To 16 $\mu$ l
GATEWAY™ BP Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 $\mu$ l	4 $\mu$ l
Total Volume	20 $\mu$ l	20 $\mu$ l	20 $\mu$ l

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80°C freezer, place immediately on ice. The mixture takes only a few minutes to thaw.

3. Add 4  $\mu$ l of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.

4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.

5. Incubate tubes at 25° for at least 60 minutes. If both the attB and attP DNAs are supercoiled, incubation for 2-24 hours at 25°C is recommended.

6. Add 2  $\mu$ l Clonase Stop Solution. Incubate for 10 min at 37°C.

7. Transform 2  $\mu$ l into 100  $\mu$ l competent E. coli, as above. Select on LB plates containing 50  $\mu$ g/ml kanamycin.

***Example 11: Cloning PCR Products into Entry Vectors using Standard Cloning Methods***

**Preparation of Entry Vectors for Cloning of PCR Products**

All of the Entry Vectors of the invention contain the death gene ccdB as a stuffer between the "left" and "right" restriction sites. The advantage of this arrangement is that there is virtually no background from vector that has not been cut with both restriction enzymes, because the presence of the ccdB gene will kill

all standard *E. coli* strains. Thus it is necessary to cut each Entry Vector twice, to remove the *ccdB* fragment.

We strongly recommend that, after digestion of the Entry Vector with the second restriction enzyme, you treat the reaction with phosphatase (calf intestine alkaline phosphatase, CIAP or thermosensitive alkaline phosphatase, TSAP). The phosphatase can be added directly to the reaction mixture, incubated for an additional time, and inactivated. This step dephosphorylates both the vector and *ccdB* fragments, so that during subsequent ligation there is less competition between the *ccdB* fragment and the DNA of interest for the termini of the Entry Vector.

10

#### Blunt Cloning of PCR products

Generally PCR products do not have 5' phosphates (because the primers are usually 5' OH), and they are not necessarily blunt. (On this latter point, see Brownstein, et al., *BioTechniques* 20: 1006, 1996 for a discussion of how the sequence of the primers affects the addition of single 3' bases.) The following protocol repairs these two defects.

15

In a 0.5 ml tube, ethanol precipitate about 40 ng of PCR product (as judged from an agarose gel).

20

1. Dissolve the precipitated DNA in 10  $\mu$ l comprising 1  $\mu$ l 10 mM rATP, 1  $\mu$ l mixed 2 mM dNTPs (i.e., 2 mM each dATP, dCTP, dTTP, and dGTP), 2  $\mu$ l 5x T4 polynucleotide kinase buffer (350 mM Tris HCl (pH7.6), 50 mM MgCl<sub>2</sub>, 500mM KCl, 5 mM 2-mercaptoethanol) 10 units T4 polynucleotide kinase, 1  $\mu$ l T4 DNA polymerase, and water to 10  $\mu$ l.
2. Incubate the tube at 37° for 10 minutes, then at 65° for 15 minutes, cool, centrifuge briefly to bring any condensate to the tip of the tube.
3. Add 5  $\mu$ l of the PEG/MgCl<sub>2</sub> solution, mix and centrifuge at room temperature for 10 minutes. Discard supernatant.
4. Dissolve the invisible precipitate in 10  $\mu$ l containing 2  $\mu$ l 5x T4 DNA ligase buffer (Life Technologies, Inc.), 0.5 units T4 DNA ligase, and about 50 ng of blunt, phosphatase-treated Entry Vector.

25

30

-120-

5. Incubate at 25° for 1 hour, then 65° for 10 minutes. Add 90 µl TE, transform 10 µl into 50 - 100 µl competent E. coli cells.
6. Plate on kanamycin.

5           **Note:** In the above protocol, steps b-c simultaneously polish the ends of the PCR product (through the exonuclease and polymerase activities of T4 DNA polymerase) and phosphorylate the 5' ends (using T4 polynucleotide kinase). It is necessary to inactivate the kinase, so that the blunt, dephosphorylated vector in step e cannot self ligate. Step d (the PEG precipitation) removes all small molecules (primers, nucleotides), and has also been found to improve the yield of cloned PCR product by 50 fold.

#### Cloning PCR Products after Digestion with Restriction Enzymes

15           Efficient cloning of PCR products that have been digested with restriction enzymes includes three steps: inactivation of *Taq* DNA polymerase, efficient restriction enzyme cutting, and removal of small DNA fragments.

20           **Inactivation of *Taq* DNA Polymerase:** Carryover of *Taq* DNA polymerase and dNTPs into a RE digestion significantly reduces the success in cloning a PCR product (D. Fox et al., *FOCUS* 20(1):15, 1998), because *Taq* DNA polymerase can fill in sticky ends and add bases to blunt ends. Either TAQQUENCH™ (obtainable from Life Technologies, Inc.; Rockville, Maryland) or extraction with phenol can be used to inactivate the *Taq*.

25           **Efficient Restriction Enzyme Cutting:** Extra bases on the 5' end of each PCR primer help the RE cut near ends of PCR products. With the availability of cheap primers, adding 6 to 9 bases on the 5' sides of the restriction sites is a good investment to ensure that most of the ends are digested. Incubation of the DNA with a 5-fold excess of restriction enzyme for an hour or more helps ensure success.

30           **Removal of Small Molecules before Ligation:** Primers, nucleotides, primer dimers, and small fragments produced by the restriction enzyme digestion,

-121-

can all inhibit or compete with the desired ligation of the PCR product to the cloning vector. This protocol uses PEG precipitation to remove small molecules.

Protocol for cutting the ends of PCR products with restriction enzyme(s):

5

1. Inactivation of Taq DNA polymerase in the PCR product:

Option A: Extraction with Phenol

10

A1. Dilute the PCR reaction to 200 µl with TE. Add an equal volume of phenol:chloroform:isoamyl alcohol, vortex vigorously for 20 seconds, and centrifuge for 1 minute at room temperature. Discard the lower phase.

15

A2. Extract the phenol from the DNA and concentrate as follows. Add an equal volume of 2-butanol (colored red with "Oil Red O" from Aldrich, if desired), vortex briefly, centrifuge briefly at room temperature. Discard the upper butanol phase. Repeat the extraction with 2-butanol. This time the volume of the lower aqueous phase should decrease significantly. Discard the upper 2-butanol phase.

20

A3. Ethanol precipitate the DNA from the aqueous phase of the above extractions. Dissolve in a 200 µl of a suitable restriction enzyme (RE) buffer.

Option B: Inactivation with TaqQuench

25

B1. Ethanol precipitate an appropriate amount of PCR product (100 ng to 1 µg), dissolve in 200 µl of a suitable RE buffer.

B2. Add 2 µl TaqQuench.

30

2. Add 10 to 50 units of restriction enzyme and incubate for at least 1 hour. Ethanol precipitate if necessary to change buffers for digestion at the other end of the PCR product.

-122-

3. Add ½ volume of the PEG/MgCl<sub>2</sub> mix to the RE digestion. Mix well and immediately centrifuge at room temperature for 10 minutes. Discard the supernatant (pellet is usually invisible), centrifuge again for a few seconds, discard any remaining supernatant.

5

4. Dissolve the DNA in a suitable volume of TE (depending on the amount of PCR product in the original amplification reaction) and apply an aliquot to an agarose gel to confirm recovery. Apply to the same gel 20-100 ng of the appropriate Entry Vector that will be used for the cloning.

10

***Example 12: Determining The Expected Size of the GATEWAY™ Cloning Reaction Products***

If you have access to a software program that will electronically cut and splice sequences, you can create electronic clones to aid you in predicting the sizes and restriction patterns of GATEWAY™ Cloning System recombination products.

The cleavage and ligation steps performed by the enzyme Int in the GATEWAY™ Cloning System recombination reactions mimic a restriction enzyme cleavage that creates a 7-bp 5'-end overhang followed by a ligation step that reseals the ends of the daughter molecules. The recombination proteins present in the Clonase cocktails (see Example 19 below) recognize the 15 bp core sequence present within all four types of att sites (in addition to other flanking sequences characteristic of each of the different types of att sites).

By treating these sites in your software program as if they were restriction sites, you can cut and splice your Entry Clones with various Destination Vectors and obtain accurate maps and sequences of the expected results from your GATEWAY™ Cloning System reactions.

30

***Example 13: Protein Expression***

**Brief Review of Protein Expression**

***Transcription:*** The most commonly used promoters in *E. coli* Expression Vectors are variants of the lac promoter, and these can be turned on by adding

IPTG to the growth medium. It is usually good to keep promoters off until expression is desired, so that the host cells are not made sick by the overabundance of some heterologous protein. This is reasonably easy in the case of the lac promoters used in *E. coli*. One needs to supply the *lac I* gene (or its more productive relative, the *lac I<sup>q</sup>* gene) to make *lac* repressor protein, which binds near the promoter and keeps transcription levels low. Some Destination Vectors for *E. coli* expression carry their own *lacI<sup>q</sup>* gene for this purpose. (However, lac promoters are always a little "on," even in the absence of IPTG.)

Controlling transcription in eukaryotic cells is not nearly so straightforward or efficient. The tetracycline system of Bujard and colleagues is the most successful approach, and one of the Destination Vectors (pDEST11; Figure 31) has been constructed to supply this function.

*Translation:* Ribosomes convert the information present in mRNA into protein. Ribosomes scan RNA molecules looking for methionine (AUG) codons, which begin nearly all nascent proteins. Ribosomes must, however, be able to distinguish between AUG codons that code for methionine in the middle of proteins from those at the start. Most often ribosomes choose AUGs that are 1) first in the RNA (toward the 5' end), and 2) have the proper sequence context. In *E. coli* the favored context (first recognized by Shine and Dalgarno, *Eur. J. Biochem.* 57: 221 (1975)) is a run of purines (As and Gs) from five to 12 bases upstream of the initiating AUG, especially AGGAGG or some variant.

In eukaryotes, a survey of translated mRNAs by Kozak (*J. Biol. Chem.* 266: 19867 (1991)) has revealed a preferred sequence context, gcc Acc ATGG, around the initiating methionine, with the A at -3 being most important, and a purine at +4 (where the A of the ATG is +1), preferably a G, being next most influential. Having an A at -3 is enough to make most ribosomes choose the first AUG of an mRNA, in plants, insects, yeast, and mammals. (For a review of initiation of protein synthesis in eukaryotic cells, see: Pain, V.M. *Eur.J. Biochem.* 236:747-771, 1996.)

*Consequences of Translation Signals for GATEWAY™ Cloning System:* First, translation signals (Shine-Dalgarno in *E. coli*, Kozak in eukaryotes) have to be close to the initiating ATG. The attB site is 25 base pairs long. Thus if

5

translation signals are desired near the natural ATG of the nucleic acid molecule of interest, they must be present in the Entry Clone of that nucleic acid molecule of interest. Also, when a nucleic acid molecule of interest is moved from an Entry Clone to a Destination vector, any translation signals will move along. The result is that the presence or absence of Shine-Dalgarno and/or Kozak sequences in the Entry Clone must be considered, with the eventual Destination Vectors to be used in mind.

10

15

20

25

Second, although ribosomes choose the 5' ATG most often, internal ATGs are also used to begin protein synthesis. The better the translation context around this internal ATG, the more internal translation initiation will be seen. This is important in the GATEWAY™ Cloning System, because you can make an Entry Clone of your nucleic acid molecule of interest, and arrange to have Shine-Dalgarno and/or Kozak sequences near the ATG. When this cassette is recombined into a Destination Vector that transcribes your nucleic acid molecule of interest, you get native protein. If you want, you can make a fusion protein in a different Destination Vector, since the Shine-Dalgarno and/or Kozak sequences do not contain any stop signals in the same reading frame. However, the presence of these internal translation signals may result in a significant amount of native protein being made, contaminating, and lowering the yield of, your fusion protein. This is especially likely with short fusion tags, like His6.

A good compromise can be recommended. If an Entry Vector like pENTR7 (Figure 16) or pENTR8 (Figure 17) is chosen, the Kozak bases are present for native eukaryotic expression. The context for *E. coli* translation is poor, so the yield of an amino-terminal fusion should be good, and the fusion protein can be digested with the TEV protease to make near-native protein following purification.

30

*Recommended Conditions for Synthesis of Proteins in E. coli:* When making proteins in *E. coli* it is advisable, at least initially, to incubate your cultures at 30°C, instead of at 37°C. Our experience indicates that proteins are less likely to form aggregates at 30°C. In addition, the yields of proteins from cells grown at 30°C frequently are improved.

5           The yields of proteins that are difficult to express may also be improved by inducing the cultures in mid-log phase of growth, using cultures begun in the morning from overnight growths, as opposed to harvesting directly from an overnight culture. In the latter case, the cells are preferably in late log or stationary growth, which can favor the formation of insoluble aggregates.

**Example 14: Constructing Destination Vectors from Existing Vectors**

10           Destination Vectors function because they have two recombination sites, attR1 and attR2, flanking a chloramphenicol resistance (CmR) gene and a death gene, ccdB. The GATEWAY™ Cloning System recombination reactions exchange the entire Cassette (except for a few bases comprising part of the attB sites) for the DNA segment of interest from the Entry Vector. Because attR1, CmR, ccdB gene, and attR2 are contiguous, they can be moved on a single DNA segment. If this Cassette is cloned into a plasmid, the plasmid becomes a  
15           Destination Vector. Figure 63 shows a schematic of the GATEWAY™ Cloning System Cassette; attR cassettes in all three reading frames contained in vectors pEZC15101, pEZC15102 and pEZC15103 are shown in Figures 64A, 64B, and 64C, respectively.

20           The protocol for constructing a Destination Vector is presented below. Keep in mind the following points:

- Destination Vectors must be constructed and propagated in one of the DB strains of *E. coli* (e.g., DB3.1, and particularly *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells) available from Life Technologies, Inc. (and described in detail in U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), because the ccdB death gene will kill any *E. coli* strain that has not been mutated such that it will survive the presence of the ccdB gene.
- If your Destination Vector will be used to make a fusion protein, a GATEWAY™ Cloning System cassette with the correct reading frame must be used. The nucleotide sequences of the ends of the cassettes are shown in Figure 78. The reading frame of the fusion protein domain must

be in frame with the core region of the attR1 site (for an amino terminal fusion) so that the six As are translated into two lysine codons. For a C-terminal fusion protein, translation through the core region of the attR2 site should be in frame with -TAC-AAA-, to yield -Tyr-Lys-.

5           • Note that each reading frame Cassette has a different unique restriction site between the chloramphenicol resistance and ccdB genes (*Mlu*I for reading frame A, *Bgl*II for reading frame B, and *Xba*I for reading frame C; see Figure 63).

10          • Most standard vectors can be converted to Destination Vectors, by inserting the Entry Cassette into the MCS of that vector.

#### Protocol for Making a Destination Vector

15          1. If the vector will make an amino fusion protein, it is necessary to keep the "aaa aaa" triplets in attR1 in phase with the triplets of the fusion protein. Determine which Entry cassette to use as follows:

20           a.) Write out the nucleotide sequence of the existing vector near the restriction site into which the Entry cassette will be cloned. These must be written in triplets corresponding to the amino acid sequence of the fusion domain.

25           b.) Draw a vertical line through the sequence that corresponds to the restriction site end, after it has been cut and made blunt, i.e., after filling in a protruding 5' end or polishing a protruding 3' end.

30           c.) Choose the appropriate reading frame cassette:

- If the coding sequence of the blunt end ends after a complete codon triplet, use the reading frame A cassette. See Figures 78, 79 and 80.

-127-

- If the coding sequence of the blunt end ends in a single base, use the reading frame B cassette. See Figures 78, 79 and 81.

5

- If the coding sequence of the blunt end ends in two bases, use the reading frame C cassette. See Figures 78, 79, 82A-B, and 83A-C.

10

2. Cut one to five micrograms of the existing plasmid at the position where you wish your nucleic acid molecule of interest (flanked by att sites) to be after the recombination reactions. **Note:** it is better to remove as many of the MCS restriction sites as possible at this step. This makes it more likely that restriction enzyme sites within the GATEWAY™ Cloning System Cassette will be unique in the new plasmid, which is important for linearizing the Destination Vector (Example 14, below).

15

3. Remove the 5' phosphates with alkaline phosphatase. While this is not mandatory, it increases the probability of success.

4. Make the end(s) blunt with fill-in or polishing reactions. For example, to 1 µg of restriction enzyme-cut, ethanol-precipitated vector DNA, add:

20

- i. 20 µl 5x T4 DNA Polymerase Buffer (165 mM Tris-acetate (pH 7.9), 330 mM Na acetate, 50 mM Mg acetate, 500 µg/ml BSA, 2.5 mM DTT)
- ii. 5 µl 10mM dNTP mix
- iii. 1 Unit of T4 DNA Polymerase
- iv. Water to a final volume of 100 µl
- v. Incubate for 15 min at 37°C.

25

30

5. Remove dNTPs and small DNA fragments: Ethanol precipitate (add three volumes of room temperature ethanol containing 0.1 M sodium acetate, mix well, immediately centrifuge at room temperature 5 - 10 minutes), dissolve wet precipitate in 200 µl TE, add 100 µl 30% PEG 8000, 30 mM MgCl<sub>2</sub>, mix well,

immediately centrifuge for 10 minutes at room temperature, discard supernatant, centrifuge again a few seconds, discard any residual liquid.

5        6. Dissolve the DNA to a final concentration of 10 - 50 ng per microliter. Apply 20 - 100 ng to a gel next to supercoiled plasmid and linear size standards to confirm cutting and recovery. The cutting does not have to be 100% complete, since you will be selecting for the chloramphenicol marker on the Entry cassette.

10        7. In a 10 µl ligation reaction combine 10 - 50 ng vector, 10 - 20 ng of Entry Cassette (Figure 79), and 0.5 units T4 DNA ligase in ligase buffer. After one hour (or overnight, whichever is most convenient), transform 1 µl into one of the DB strains of competent *E. coli* cells with a *gyrA*462 mutation (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), preferably DB3.1, and most preferably *E. coli* LIBRARY 15 EFFICIENCY® DB3.1™ Competent Cells. The *ccdB* gene on the Entry Cassette will kill other strains of *E. coli* that have not been mutated so as to survive the presence of the *ccdB* gene.

20        8. After expression in SOC medium, plate 10 µl and 100 µl on chloramphenicol-containing (30 µg / ml) plates, incubate at 37° C.

25        9. Pick colonies, make miniprep DNA. Treat the miniprep with RNase A and store in TE. Cut with the appropriate restriction enzyme to determine the orientation of the Cassette. Choose clones with the attR1 site next to the amino end of the protein expression function of the plasmid.

#### Notes on Using Destination Vectors

- We have found that about ten-fold more colonies result from a GATEWAY™ Cloning System reaction if the Destination Vector is linear or relaxed. If the competent cells you use are highly competent (>10<sup>8</sup> per microgram), linearizing the Destination Vector is less essential.

-129-

- The site or sites used for the linearization must be within the Entry Cassette. Sites that cut once or twice within each cassette are shown in Figures 80-82.
- Minipreps of Destination Vectors will work fine, so long as they have been treated with RNase. Since most DB strains are *endA*- (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), minipreps can be digested with restriction enzymes without a prior phenol extraction.
- Reading the OD<sub>260</sub> of miniprep DNA is inaccurate unless the RNA and ribonucleotides have been removed, for example, by a PEG precipitation.

5

10

***Example 15: Some Options in Choosing Appropriate Entry Vectors and Destination Vectors: An Example***

15

In some applications, it may be desirable to express a nucleic acid molecule of interest in two forms: as an amino-terminal fusion in *E. coli*, and as a native protein in eukaryotic cells. This may be accomplished in any of several ways:

20

**Option 1:** Your choices depend on your nucleic acid molecule of interest and the fragment that contains it, as well as the available Entry Vectors. For eukaryotic translation, you need consensus bases according to Kozak (*J. Biol. Chem.* 266:19867, 1991) near the initiating methionine (ATG) codon. All of the Entry Vectors offer this motif upstream of the *XmnI* site (blunt cutter). One option is to amplify your nucleic acid molecule of interest, with its ATG, by PCR, making the amino end blunt and the carboxy end containing the natural stop codon followed by one of the "right side" restriction sites (*EcoRI*, *NotI*, *XhoI*, *EcoRV*, or *XbaI* of the pENTR vectors).

25

If you know your nucleic acid molecule of interest does not have, for example, an *XhoI* site, you can make a PCR product that has this structure:

30

*Xho I*

5' ATG nnn nnn --- nnn TAA ctc gag nnn nnn 3'  
3' tac nnn nnn --- nnn att gag ctc nnn nnn 5'

-130-

After cutting with *Xhol*, the fragment is ready to clone:

5' ATG nnn nnn --- nnn TAA c            3'  
3' tac nnn nnn --- nnn att gag ct    5'

5 (If you follow this example, don't forget to put a phosphate on the amino oligo.)

10            **Option 2:** This PCR product could be cloned into two Entry Vectors to give the desired products, between the *XmnI* and *XhoI* sites: pENTR1A (Figures 10A, 10B) or pENTR7 (Figures 16A, 16B). If you clone into pENTR1A, amino fusions will have the minimal number of amino acids between the fusion domain and your nucleic acid molecule of interest, but the fusion cannot be removed with TEV protease. The converse is true of clones in pENTR7, i.e., an amino fusion can be cleaved with TEV protease, at the cost of more amino acids between the fusion and your nucleic acid molecule of interest.

15            In this example, let us choose to clone our hypothetical nucleic acid molecule of interest into pENTR7, between the *XmnI* and *XhoI* sites. Once this is accomplished, several optional protocols using the Entry Clone pENTR7 may be followed:

20            **Option 3:** Since the nucleic acid molecule of interest has been amplified with PCR, it may be desirable to sequence it. To do this, transfer the nucleic acid molecule of interest from the Entry Vector into a vector that has priming sites for the standard sequencing primers. Such a vector is pDEST6 (Figures 26A, 26B). This Destination Vector places the nucleic acid molecule of interest in the opposite orientation to the lac promoter (which is leaky -- see Example 3 above). If the gene product is toxic to *E. coli*, this Destination Vector will minimize its toxicity.

25            **Option 4:** While the sequencing is going on, you might wish to check the expression of the nucleic acid molecule of interest in, for example, CHO cells, by recombining the nucleic acid molecule of interest into a CMV promoter vector (pDEST7, Figure 27; or pDEST12, Figure 32), or into a baculovirus vector (pDEST8, Figure 28; or pDEST10, Figure 30) for expression in insect cells. Both

-131-

of these vectors will transcribe the coding sequence of your nucleic acid molecule of interest, and translate it from the ATG of the PCR product using the Kozak bases upstream of the *XmnI* site.

5           **Option 5:** If you wish to purify protein, for example to make antibodies, you can clone the nucleic acid molecule of interest into a His6 fusion vector, pDEST2 (Figure 22). Since the nucleic acid molecule of interest is cloned downstream of the TEV protease cleavage domain of pENTR7 (Figure 16), the amino acid sequence of the protein produced will be:

10           [----- attB1 -----]           TEV protease  
NH2- MSYYHHHHHHGITSLYKKAGF*ENLYFQ!* GTM---COOH

15           The attB site and the restriction sites used to make the Destination and Entry Vectors are translated into the underlined 11 amino acids (GITSLYKKAGF). Cleavage with TEV protease (arrow) leaves two amino acids, 20           GT, on the amino end of the gene product.

See Figure 55 for an example of a nucleic acid molecule of interest, the chloramphenicol acetyl transferase (CAT) gene, cloned into pENTR7 (Figure 16) as a blunt (amino)-*Xba*I (carboxy) fragment, then cloned by recombination into the His6 fusion vector pDEST2 (Figure 22).

25           **Option 6:** If the His6 fusion protein is insoluble, you may go on and try a GST fusion. The appropriate Destination vector is pDEST3 (Figure 23).

30           **Option 7:** If you need to make RNA probes and prefer SP6 RNA polymerase, you can make the top strand RNA with your nucleic acid molecule of interest cloned into pSPORT+ (pDEST5 (Figures 25A, 25B)), and the bottom strand RNA with the nucleic acid molecule of interest cloned into pSPORT(-) (pDEST6 (Figures 26A, 26B)). Opposing promoters for T7 RNA polymerase and SP6 RNA polymerase are also present in these clones.

5                   **Option 8:** It is often worthwhile to clone your nucleic acid molecule of interest into a variety of Destination Vectors in the same experiment. For example, if the number of colonies varies widely when the various recombination reactions are transformed into *E. coli*, this may be an indication that the nucleic acid molecule of interest is toxic in some contexts. (This problem is more clearly evident when a positive control gene is used for each Destination Vector.) Specifically, if many more colonies are obtained when the nucleic acid molecule of interest is recombined into pDEST6 than in pDEST5, there is a good chance that leakiness of the lac promoter is causing some expression of the nucleic acid molecule of interest in pSPORT "+" (which is not harmful in pDEST6 because the nucleic acid molecule of interest is in the opposite orientation).

10

15                  *Example 16: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction*

20                  In the BxP recombination (Entry or Gateward) reaction described herein, a DNA segment flanked by attB1 and attB2 sites in a plasmid conferring ampicillin resistance was transferred by recombination into an attP plasmid conferring kanamycin resistance, which resulted in a product molecule wherein the DNA segment was flanked by attL sites (attL1 and attL2). This product plasmid comprises an "attL Entry Clone" molecule, because it can react with a "attR Destination Vector" molecule via the LxR (Destination) reaction, resulting in the transfer of the DNA segment to a new (ampicillin resistant) vector. In the previously described examples, it was necessary to transform the BxP reaction products into *E. coli*, select kanamycin resistant colonies, grow those colonies in liquid culture, and prepare miniprep DNA, before reacting this DNA with a Destination Vector in an LxR reaction.

25

30                  The goal of the following experiment was to eliminate the transformation and miniprep DNA steps, by adding the BxP Reaction products directly to an LxR Reaction. This is especially appropriate when the DNA segment flanked by attB sites is a PCR product instead of a plasmid, because the PCR product cannot give

-133-

ampicillin-resistant colonies upon transformation, whereas attB plasmids (in general) carry an ampicillin resistance gene. Thus use of a PCR product flanked by attB sites in a BxP Reaction allows one to select for the ampicillin resistance encoded by the desired attB product of a subsequent LxR Reaction.

5 Two reactions were prepared: Reaction A, negative control, no attB PCR product, (8 µl) contained 50 ng pEZC7102 (attP Donor plasmid, confers kanamycin resistance) and 2 µl BxP Clonase (22 ng / µl Int protein and 8 ng/µl IHF protein) in BxP buffer (25 mM Tris HCl, pH 7.8, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, 250 µg / ml BSA). Reaction B (24 µl) contained  
10 150 ng pEZC7102, 6 µl BxP Clonase, and 120 ng of the attB -tet-PCR product in the same buffer as reaction A. The attB - tet - PCR product comprised the tetracycline resistance gene of plasmid pBR322, amplified with two primers containing either attB1 or attB2 sites, and having 4 Gs at their 5' ends, as described earlier.

15 The two reactions were incubated at 25°C for 30 minutes. Then aliquots of these reactions were added to new components that comprised LxR Reactions or appropriate controls for the LxR Reaction. Five new reactions were thus produced:

20 **Reaction 1:** 5 µl of reaction A was added to a 5 µl LxR Reaction containing 25 ng *NcoI*-cut pEZC8402 (the attR Destination Vector plasmid) in LxR buffer (37.5 mM Tris HCl, pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA), and 1 µl of GATEWAY™ LR Clonase™ Enzyme Mix (total volume of 10 µl).

25 **Reaction 2:** Same as reaction 1, except 5 µl of reaction B (positive) were added instead of reaction A (negative).

30 **Reaction 3:** Same as reaction 2, except that the amounts of Nco-cut pEZC8402 and GATEWAY™ LR Clonase™ Enzyme Mix were doubled, to 50 ng and 2 µl, respectively.

-134-

**Reaction 4:** Same as reaction 2, except that 25 ng of pEZ11104 (a positive control attL Entry Clone plasmid) were added in addition to the aliquot of reaction B.

5           **Reaction 5:** Positive control LxR Reaction, containing 25 ng *Nco*I-cut pEZZC8402, 25 ng pEZ11104, 37.5 mM Tris HCl pH 7.7, 16.5 mM NaCl, 35 mM KCl; 5 mM spermidine, 375 µg / ml BSA and 1 µl GATEWAY™ LR Clonase™ Enzyme Mix in a total volume of 5 µl.

10           All five reactions were incubated at 25°C for 30 minutes. Then, 1 µl aliquots of each of the above five reactions, plus 1 µl from the remaining volume of Reaction B, the standard BxP Reaction, were used to transform 50 µl competent DH5 $\alpha$  *E. coli*. DNA and cells were incubated on ice for 15 min., heat shocked at 42°C for 45 sec., and 450 µl SOC were added. Each tube was incubated with shaking at 37°C for 60 min. Aliquots of 100 µl and 400 µl of each transformation were plated on LB plates containing either 50 µg/ml kanamycin or 100 µg/ml ampicillin (see Table 2). A transformation with 10 pg of pUC19 DNA (plated on LB-amp<sub>100</sub>) served as a control on the transformation efficiency of the DH5 $\alpha$  cells. Following incubation overnight at 37°C, the number of colonies on each plate was determined.

15           20           Results of these reactions are shown in Table 2.

Table 2\*

Reaction No.:	1	2	3	4	5	6
Number of Colonies						
Vol. plated: BxP Reaction	Neg Control pEZZC8402 and LR Clonase™	1X pEZZC8402 and LR Clonase™	2X pEZZC8402 and LR Clonase™	LxR Reaction with Pos. Control DNA	LxR Reaction alone	BxP Reaction alone
100 µl	2	1	8	9	~1000	~1000
400 µl	5	10	35	62	>2000	>2000
Selection:	Kan	Amp	Amp	Amp	Amp	Kan

\*(Transformation with pUC 19 DNA yielded 1.4 x 10<sup>9</sup> CFU/µg DNA.)

34 of the 43 colonies obtained from Reaction 3 were picked into 2 ml Terrific Broth with 100 µg/ml ampicillin and these cultures were grown overnight, with shaking, at 37°C. 27 of the 34 cultures gave at least moderate growth, and of these 24 were used to prepare miniprep DNA, using the standard protocol.

5 These 24 DNAs were initially analyzed as supercoiled (SC) DNA on a 1% agarose gel to identify those with inserts and to estimate the sizes of the inserts. Fifteen of the 24 samples displayed SC DNA of the size predicted (5553 bp) if **tetx7102** had correctly recombined with **pEYC8402** to yield **tetx8402**. One of these samples contained two plasmids, one of ~5500 bp and a one of ~3500 bp. The 10 majority of the remaining clones were approximately 4100 bp in size.

All 15 of the clones displaying SC DNA of predicted size (~5500 bp) were analyzed by two different double digests with restriction endonucleases to confirm the structure of the expected product: **tetx8402**. (See plasmid maps, Figures 57-59) In one set of digests, the DNAs were treated with *Not I* and *Eco RI*, which should cut the predicted product just outside both *attB* sites, releasing the *tet<sup>r</sup>* insert on a fragment of 1475 bp. In the second set of digests, the DNAs were digested with *NotI* and with *NruI*. *NruI* cleaves asymmetrically within the subcloned *tet<sup>r</sup>* insert, and together with *NotI* will release a fragment of 1019 bp.

15 Of the 15 clones analyzed by double restriction digestion, 14 revealed the 20 predicted sizes of fragments for the expected product.

### Interpretation:

The DNA components of Reaction B, pEYC7102 and *attB-tet-PCR*, are shown in Figure 56. The desired product of BxP Reaction B is **tetx7102**, depicted in Figure 57. The LxR Reaction recombines the product of the BxP Reaction, **tetx7102** (Figure 57), with the Destination Vector, **pEYC8402**, shown in Figure 58. The LxR Reaction with **tetx7102** plus **pEYC8402** is predicted to yield the desired product **tetx8402**, shown in Figure 59.

30 Reaction 2, which combined the BxP Reaction and LxR Reaction, gave few colonies beyond those of the negative control Reaction. In contrast, Reaction 3, with twice the amount of **pEYC8402** (Figure 58) and LxR Clonase, yielded a

larger number of colonies. These colonies were analyzed further, by restriction digestion, to confirm the presence of expected product. Reaction 4 included a known amount of attL Entry Clone plasmid in the combined BxP-plus-LxR reaction. But reaction 4 yielded only about 1% of the colonies obtained when the same DNA was used in a LxR reaction alone, Reaction 6. This result suggests 5 that the LxR reaction may be inhibited by components of the BxP reaction.

Restriction endonuclease analysis of the products of Reaction 3 revealed that a sizeable proportion of the colonies (14 of the 34 analyzed) contained the desired tet<sup>r</sup> subclone, tetx8402 (Figure 59).

10 The above results establish the feasibility of performing first a BxP recombination reaction followed by a LxR recombination reaction -- in the same tube -- simply by adding the appropriate buffer mix, recombination proteins, and DNAs to a completed BxP reaction. This method should prove useful as a faster method to convert attB-containing PCR products into different Expression 15 Clones, eliminating the need to isolate first the intermediate attL-PCR insert subclones, before recombining these with Destination Vectors. This may prove especially valuable for automated applications of these reactions.

This same one-tube approach allows for the rapid transfer of nucleic acid 20 molecules contained in attB plasmid clones into new functional vectors as well. As in the above examples, attL subclones generated in a BxP Reaction can be recombined directly with various Destination Vectors in a LxR reaction. The only additional requirement for using attB plasmids, instead of attB-containing PCR products, is that the Destination Vector(s) employed must contain a different selection marker from the one present on the attB plasmid itself and the attP vector. 25

Two alternative protocols for a one-tube reaction have also proven useful and somewhat more optimal than the conditions described above.

Alternative 1:

30 Reaction buffer contained 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.25 mM EDTA, 2.5 mM spermidine, and 200 µg/ml BSA. After a 16 (or 3) hour incubation of the PCR product (100 ng) + attP Donor plasmid (100 ng) +

-137-

GATEWAY™ BP Clonase™ Enzyme Mix + Destination Vector (100 ng), 2 µl of GATEWAY™ LR Clonase™ Enzyme Mix (per 10 µl reaction mix) was added and the mixture was incubated an additional 6 (or 2) hours at 25°C. Stop solution was then added as above and the mixture was incubated at 37°C as above and transformed by electroporation with 1 µl directly into electrocompetent host cells. Results of this series of experiments demonstrated that longer incubation times (16 hours vs. 3 hours for the BP Reaction, 6 hours vs. 2 hours for the LR Reaction) resulted in about twice as many colonies being obtained as for the shorter incubation times. With two independent genes, 10/10 colonies having the correct cloning patterns were obtained.

5

10

15

20

25

30

Alternative 2:

A standard BP Reaction under the reaction conditions described above for Alternative 1 was performed for 2 hours at 25°C. Following the BP Reaction, the following components were added to the reaction mixture in a total volume of 7 µl:

20 mM Tris-HCl, pH 7.5

100 mM NaCl

5 µg/ml Xis-His6

15% glycerol

~1000 ng of Destination Vector

The reaction mixture was then incubated for 2 hours at 25°C, and 2.5 µl of stop solution (containing 2 µg/ml proteinase K) was added and the mixture was incubated at 37°C for an additional 10 minutes. Chemically competent host cells were then transformed with 2 µl of the reaction mixture, or electrocompetent host cells (*e.g.*, EMax DH10B cells; Life Technologies, Inc.) were electroporated with 2 µl of the reaction mixture per 25-40 µl of cells. Following transformation, mixtures were diluted with SOC, incubated at 37°C, and plated as described above on media selecting for the selection markers on the Destination Vector and the Entry clone (B x P reaction product). Analogous results to those described for Alternative 1 were obtained with these reaction conditions -- a higher level of colonies containing correctly recombined reaction products were observed.

**Example 17: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction**

5 Single-tube transfer of PCR product DNA or Expression Clones into Expression Clones by recombinational cloning has also been accomplished using a procedure modified from that described in Example 16. This procedure is as follows:

10 • Perform a standard BP (Gateward) Reaction (see Examples 9 and 10) in 20 µl volume at 25°C for 1 hour.

15 • After the incubation is over, take a 10 µl aliquot from the 20 µl total volume and add 1 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes. This first aliquot can be used for transformation and gel assay of BP reaction analysis. Plate BP reaction transformation on LB plates with **Kanamycin** (50 µg/ml).

20 • Add the following reagents to the remaining 10 µl aliquot of the BP reaction:

1 µl of 0.75 M NaCl

2 µl of destination vector (150 ng/µl)

4 µl of LR Clonase™ (after thawing and brief mixing)

25 • Mix all reagents well and incubate at 25°C for 3 hours. Stop the reaction at the end of incubation with 1.7 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes.

30 • Transform 2 µl of the completed reaction into 100 µl of competent cells. Plate 100 µl and 400 µl on LB plates with **Ampicillin** (100 µg/ml).

**Notes:**

- If your competent cells are less than 10<sup>8</sup> CFU/µg, and you are concerned about getting enough colonies, you can improve the yield several fold by incubating the

-139-

BP reaction for 6-20 hours. Electroporation also can yield better colony output than chemical transformation.

5 •PCR products greater than about 5-6 kb show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using longer incubation times for both BP and LR steps.

10 •If you want to move your insert gene into several destination vectors simultaneously, then scale up the initial BP reaction volume so that you have a 10 µl aliquot for adding each destination vector.

***Example 18: Optimization of GATEWAY™ Clonase™ Enzyme Compositions***

15 The enzyme compositions containing Int and IHF (for BP Reactions) were optimized using a standard functional recombinational cloning reaction (a BP reaction) between attB-containing plasmids and attP-containing plasmids, according to the following protocol:

**Materials and Methods:**

20 ***Substrates:***

AttP - supercoiled pDONR201

AttB - linear ~1Kb [<sup>3</sup>H]PCR product amplified from pEZC7501

25 ***Proteins:***

IntH6 -- His<sub>6</sub>-carboxy- tagged λ Integrase

IHF -- Integration Host Factor

***Clonase:***

30 50 ng/µl IntH6 and 20 ng/µl IHF, admixed in 25 mM Tris- HCl (pH 7.5), 22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, and 50% glycerol.

-140-

*Reaction Mixture (total volume of 40 µl):*

1000 ng AttP plasmid

600 ng AttB [<sup>3</sup>H] PCR product

8 µl Clonase (400 ng IntH6, 160 ng IHF) in 25 mM Tris-HCl (pH 7.5),  
5  
22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, 5 mM  
DTT.

Reaction mixture was incubated for 1 hour at 25°C, 4 µl of 2 µg/µl  
10 proteinase K was added and mixture was incubated for an additional 20 minutes  
at 37°C. Mixture was then extracted with an equal volume of Phenol/Chloroform/  
Isoamyl alcohol. The aqueous layer was then collected, and 0.1 volumes of 3 M  
15 sodium acetate and 2 volumes of cold 100% ethanol were added. Tubes were  
then spun in a microcentrifuge at maximum RPM for 10 minutes at room  
temperature. Ethanol was decanted, and pellets were rinsed with 70% ethanol and  
re-centrifuged as above. Ethanol was decanted, and pellets were allowed to air  
dry for 5-10 minutes and then dissolved in 20 µl of 33 mM Tris-Acetate (pH 7.8),  
66 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, and 1mM  
20 ATP. 2 units of exonuclease V (e.g., Plasmid Safe; EpiCentre, Inc., Madison, WI)  
was then added, and the mixture was incubated at 37°C for 30 minutes.

Samples were then TCA-washed by spotting 30 µl of reaction mixture  
onto a Whatman GF/C filter, washing filters once with 10% TCA + 1% NaPPi for  
10 minutes, three times with 5% TCA for 5 minutes each, and twice with ethanol  
for 5 minutes each. Filters were then dried under a heat lamp, placed into a  
25 scintillation vial, and counted on a β liquid scintillation counter (LSC).

The principle behind this assay is that, after exonuclease V digestion, only  
30 double-stranded circular DNA survives in an acid-insoluble form. All DNA  
substrates and products that have free ends are digested to an acid-soluble form  
and are not retained on the filters. Therefore, only the <sup>3</sup>H-labeled attB linear DNA  
which ends up in circular form after both inter- and intramolecular integration is  
complete is resistant to digestion and is recovered as acid-insoluble product.  
Optimal enzyme and buffer formulations in the Clonase compositions therefore are  
those that give the highest levels of circularized <sup>3</sup>H-labeled attB-containing

sequences, as determined by highest cpm in the LSC. Although this assay was designed for optimization of GATEWAY™ BP Clonase™ Enzyme Mix compositions (Int + IHF), the same type of assay may be performed to optimize GATEWAY™ LR Clonase™ Enzyme Mix compositions (Int + IHF + Xis), except that the reaction mixtures would comprise 1000 ng of AttR (instead of AttP) and 600 ng of AttL (instead of AttB), and 40 ng of His<sub>6</sub>-carboxy- tagged Xis (XisH6) in addition to the IntH6 and IHF.

***Example 19: Testing Functionality of Entry and Destination Vectors***

As part of assessment of the functionality of particular vectors of the invention, it is important to functionally test the ability of the vectors to recombine. This assessment can be carried out by performing a recombinational cloning reaction (as schematized in Figures 2, 4, and 5A and 5B, and as described herein and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of all of which are incorporated by reference herein in their entireties), by transforming *E. coli* and scoring colony forming units. However, an alternative assay may also be performed to allow faster, more simple assessment of the functionality of a given Entry or Destination Vector by agarose gel electrophoresis. The following is a description of such an in vitro assay.

**Materials and Methods:**

Plasmid templates pEYC1301 (Figure 84) and pEYC1313 (Figure 85), each containing a single wild type att site, were used for the generation of PCR products containing attL or attR sites, respectively. Plasmid templates were linearized with *Afl*NI, phenol extracted, ethanol precipitated and dissolved in TE to a concentration of 1 ng/ $\mu$ l.

-142-

PCR primers (capital letters represent base changes from wildtype):

attL1	gggg agcct gc <sup>T</sup> ttttGtacAaa gttggcatta taaaaaagca ttgc
attL2	gggg agcct gc <sup>T</sup> ttCttGtacAaa gttggcatta taaaaaagca ttgc
attL right	tgttgccggg aagctagagt aa
5	
attR1	gggg Acaag ttTgtAaaaaaagc tgaacgaga aacgtaaaat
attR2	gggg Acaag ttTgtAaaGaaagc tgaacgaga aacgtaaaat
attR right	ca gacggcatga tgaacctgaa

10 PCR primers were dissolved in TE to a concentration of 500 pmol/μl. Primer mixes were prepared, consisting of attL1 + attLright primers, attL2 + attLright primers, attR1 + attRight primers, and attR2 + attRight primers, each mix containing 20 pmol/μl of each primer.

PCR reactions:

15 1 μl plasmid template (1 ng)  
1 μl primer pairs (20 pmoles of each)  
3 μl of H<sub>2</sub>O  
45 μl of Platinum PCR SuperMix® (Life Technologies, Inc.)

20 Cycling conditions (performed in MJ thermocycler):

95°C/2 minutes  
94°C/30 seconds  
25 cycles of 58°C/30 seconds and 72°C/1.5 minutes  
72°C/5 minutes  
5°C/hold

25

The resulting attL PCR product was 1.5 kb, and the resulting attR PCR product was 1.0 kb.

30 PCR reactions were PEG/MgCl<sub>2</sub> precipitated by adding 150 μl H<sub>2</sub>O and 100 μl of 3x PEG/ MgCl<sub>2</sub> solution followed by centrifugation. The PCR products were dissolved in 50 μl of TE. Quantification of the PCR product was performed by gel electrophoresis of 1 μl and was estimated to be 50-100 ng/μl.

Recombination reactions of PCR products containing attL or attR sites with GATEWAY™ plasmids was performed as follows:

8 µl of H<sub>2</sub>O  
2 µl of attL or attR PCR product (100-200 ng)  
5 2 µl of GATEWAY™ plasmid (100 ng)  
4 µl of 5x Destination buffer  
4 µl of GATEWAY™ LR Clonase™ Enzyme Mix

20 µl total volume (the reactions can be scaled down to a 5 µl total volume by adjusting the volumes of the components to about ¼ of those shown above, while keeping the stoichiometries the same).

Clonase reactions were incubated at 25 °C for 2 hours. 2 µl of proteinase K (2 mg/ml) was added to stop the reaction. 10 µl was then run on a 1 % agarose gel. Positive control reactions were performed by reacting attL1 PCR product (1.0 kb) with attR1 PCR product (1.5 kb) and by similarly reacting attL2 PCR product with attR2 PCR product to observe the formation of a larger (2.5 kb) recombination product. Negative controls were similarly performed by reacting attL1 PCR product with attR2 PCR product and vice versa or reactions of attL PCR product with an attL plasmid, etc.

In alternative assays, to test attB Entry vectors, plasmids containing single attP sites were used. Plasmids containing single att sites could also be used as recombination substrates in general to test all Entry and Destination vectors (*i.e.*, those containing attL, attR, attB and attP sites). This would eliminate the need to do PCR reactions.

25 Results:

Destination and Entry plasmids when reacted with appropriate att-containing PCR products formed linear recombinant molecules that could be easily visualized on an agarose gel when compared to control reactions containing no attL or attR PCR product. Thus, the functionality of Destination and Entry vectors constructed according to the invention may be determined either by carrying out the Destination or Entry recombination reactions as depicted in

-144-

Figures 2, 4, and 5A and 5B, or more rapidly by carrying out the linearization assay described in this Example.

***Example 20: PCR Cloning Using Universal Adapter-Primers***

As described herein, the cloning of PCR products using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) requires the addition of attB sites (attB1 and attB2) to the ends of gene-specific primers used in the PCR reaction. The protocols described in the preceding Examples suggest that the user add 29 bp (25 bp containing the attB site plus four G residues) to the gene-specific primer. It would be advantageous to high volume users of the GATEWAY™ PCR Cloning System to generate attB-containing PCR product using universal attB adapter-primers in combination with shorter gene-specific primers containing a specified overlap to the adapters. The following experiments demonstrate the utility of this strategy using universal attB adapter-primers and gene-specific primers containing overlaps of various lengths from 6 bp to 18 bp. The results demonstrate that gene-specific primers with overlaps of 10 bp to 18 bp can be used successfully in PCR amplifications with universal attB adapter-primers to generate full-length PCR products. These PCR products can then be successfully cloned with high fidelity in a specified orientation using the GATEWAY™ PCR Cloning System.

**Methods and Results:**

To demonstrate that universal attB adapter-primers can be used with gene-specific primers containing partial attB sites in PCR reactions to generate full-length PCR product, a small 256 bp region of the human hemoglobin cDNA was chosen as a target so that intermediate sized products could be distinguished from full-length products by agarose gel electrophoresis.

The following oligonucleotides were used:

B1-Hgb: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-5'-Hgb\*  
B2-Hgb: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T-3'-Hgb\*\*

-145-

18B1-Hgb:	TG TAC AAA AAA GCA GGC T-5'-Hgb
18B2-Hgb:	TG TAC AAG AAA GCT GGG T-3'-Hgb
15B1-Hgb:	AC AAA AAA GCA GGC T-5'-Hgb
15B2-Hgb:	AC AAG AAA GCT GGG T-3'-Hgb
5 12B1-Hgb:	AA AAA GCA GGC T-5'-Hgb
12B2-Hgb:	AG AAA GCT GGG T-3'-Hgb
11B1-Hgb:	A AAA GCA GGC T-5'-Hgb
11B2-Hgb:	G AAA GCT GGG T-3'-Hgb
10 10B1-Hgb:	AAA GCA GGC T-5'-Hgb
10B2-Hgb:	AAA GCT GGG T-3'-Hgb
9B1-Hgb:	AA GCA GGC T-5'-Hgb
9B2-Hgb:	AA GCT GGG T-3'-Hgb
8B1-Hgb:	A GCA GGC T-5'-Hgb
8B2-Hgb:	A GCT GGG T-3'-Hgb
15 7B1-Hgb:	GCA GGC T-5'-Hgb
7B2-Hgb:	GCT GGG T-3'-Hgb
6B1-Hgb:	CA GGC T-5'-Hgb
6B2-Hgb:	CT GGG T-3'-Hgb

20 attB1 adapter: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T  
attB2 adapter: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T

\* -5'-Hgb = GTC ACT AGC CTG TGG AGC AAG A

\*\* -3'-Hgb = AGG ATG GCA GAG GGA GAC GAC A

25

The aim of these experiments was to develop a simple and efficient universal adapter PCR method to generate attB containing PCR products suitable for use in the GATEWAY™ PCR Cloning System. The reaction mixtures and thermocycling conditions should be simple and efficient so that the universal adapter PCR method could be routinely applicable to any PCR product cloning application.

30

PCR reaction conditions were initially found that could successfully amplify predominately full-length PCR product using gene-specific primers containing 18bp and 15 bp overlap with universal attB primers. These conditions are outlined below:

35

-146-

10 pmoles of gene-specific primers  
10 pmoles of universal attB adapter-primers  
1 ng of plasmid containing the human hemoglobin cDNA.  
100 ng of human leukocyte cDNA library DNA.  
5  $\mu$ l of 10x PLATINUM Taq HiFi® reaction buffer (Life Technologies, Inc.)  
2  $\mu$ l of 50 mM MgSO<sub>4</sub>  
1  $\mu$ l of 10 mM dNTPs  
0.2  $\mu$ l of PLATINUM Taq HiFi® (1.0 unit)  
H<sub>2</sub>O to 50  $\mu$ l total reaction volume

10

Cycling conditions:

15

25 x | 95°C/5 min  
          | 94°C/15 sec  
          | 50°C/30 sec  
          | 68°C/1 min  
          | 68°C/5 min  
          | 5°C/hold

20 To assess the efficiency of the method, 2  $\mu$ l (1/25) of the 50  $\mu$ l PCR reaction was electrophoresed in a 3 % Agarose-1000 gel. With overlaps of 12 bp or less, smaller intermediate products containing one or no universal attB adapter predominated the reactions. Further optimization of PCR reaction conditions was obtained by titrating the amounts of gene-specific primers and universal attB adapter-primers. The PCR reactions were set up as outlined above except that the 25 amounts of primers added were:

0, 1, 3 or 10 pmoles of gene-specific primers  
0, 10, 30 or 100 pmoles of adapter-primers

-147-

## Cycling conditions:

	95°C/3 min
	94°C/15 sec
25 x	50°C/45 sec
	68°C/1 min
	68°C/5 min
	5°C/hold

5

The use of limiting amounts of gene-specific primers (3 pmoles) and excess adapter-primers (30 pmoles) reduced the amounts of smaller intermediate products. Using these reaction conditions the overlap necessary to obtain predominately full-length PCR product was reduced to 12 bp. The amounts of gene-specific and adapter-primers was further optimized in the following PCR reactions:

10

0, 1, 2 or 3 pmoles of gene-specific primers  
0, 30, 40 or 50 pmoles of adapter-primers

15

## Cycling conditions:

	95°C/3 min
	94°C/15 sec
25 x	48°C/1 min
	68°C/1 min
	68°C/5 min
25	5°C/hold

20

25

The use of 2 pmoles of gene-specific primers and 40 pmoles of adapter-primers further reduced the amounts of intermediate products and generated predominately full-length PCR products with gene-specific primers containing an 11 bp overlap. The success of the PCR reactions can be assessed in any PCR application by performing a no adapter control. The use of limiting amounts of gene-specific primers should give faint or barely visible bands when 1/25 to 1/10 of the PCR reaction is electrophoresed on a standard agarose gel. Addition of the

30

universal attB adapter-primers should generate a robust PCR reaction with a much higher overall yield of product.

PCR products from reactions using the 18 bp, 15 bp, 12 bp, 11 bp and 10 bp overlap gene-specific primers were purified using the CONCERT® Rapid PCR Purification System (PCR products greater than 500 bp can be PEG precipitated). The purified PCR products were subsequently cloned into an attP containing plasmid vector using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) and transformed into *E. coli*. Colonies were selected and counted on the appropriate antibiotic media and screened by PCR for correct inserts and orientation.

Raw PCR products (unpurified) from the attB adapter PCR of a plasmid clone of part of the human beta-globin (Hgb) gene were also used in GATEWAY™ PCR Cloning System reactions. PCR products generated with the full attB B1/B2-Hgb, the 12B1/B2, 11B1/B2 and 10B1/B2 attB overlap Hgb primers were successfully cloned into the GATEWAY™ pENTR21 attP vector (Figure 49). 24 colonies from each (24 x 4 = 96 total) were tested and each was verified by PCR to contain correct inserts. The cloning efficiency expressed as cfu/ml is shown below:

Primer Used	cfu/ml
Hgb full attB	8,700
Hgb 12 bp overlap	21,000
Hgb 11 bp overlap	20,500
Hgb 10 bp overlap	13,500
GFP control	1,300

Interestingly, the overlap PCR products cloned with higher efficiency than did the full attB PCR product. Presumably, and as verified by visualization on agarose gel, the adapter PCR products were slightly cleaner than was the full attB PCR product. The differences in colony output may also reflect the proportion of PCR product molecules with intact attB sites.

Using the attB adapter PCR method, PCR primers with 12 bp attB overlaps were used to amplify cDNAs of different sizes (ranging from 1 to 4 kb)

from a leukocyte cDNA library and from first strand cDNA prepared from HeLa total RNA. While three of the four cDNAs were able to be amplified by this method, a non-specific amplification product was also observed that under some conditions would interfere with the gene-specific amplification. This non-specific product was amplified in reactions containing the attB adapter-primers alone without any gene-specific overlap primers present. The non-specific amplification product was reduced by increasing the stringency of the PCR reaction and lowering the attB adapter PCR primer concentration.

These results indicate that the adapter-primer PCR approach described in this Example will work well for cloned genes. These results also demonstrate the development of a simple and efficient method to amplify PCR products that are compatible with the GATEWAY™ PCR Cloning System that allows the use of shorter gene-specific primers that partially overlap universal attB adapter-primers. In routine PCR cloning applications, the use of 12 bp overlaps is recommended. The methods described in this Example can thus reduce the length of gene-specific primers by up to 17 residues or more, resulting in a significant savings in oligonucleotide costs for high volume users of the GATEWAY™ PCR Cloning System. In addition, using the methods and assays described in this Example, one of ordinary skill can, using only routine experimentation, design and use analogous primer-adapters based on or containing other recombination sites or fragments thereof, such as *attL*, *attR*, *attP*, *lox*, FRT, etc.

**Example 21: Mutational Analysis of the Bacteriophage Lambda *attL* and *attR* Sites: Determinants of *att* Site Specificity in Site-specific Recombination**

To investigate the determinants of *att* site specificity, the bacteriophage lambda *attL* and *attR* sites were systematically mutagenized. As noted herein, the determinants of specificity have previously been localized to the 7 bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four lambda *att* sites, *attB*, *attP*, *attL* and *attR*. This core region, however, has not heretofore been systematically

mutagenized and examined to define precisely which mutations produce unique changes in *att* site specificity.

Therefore, to examine the effect of *att* sequence on site specificity, mutant *attL* and *attR* sites were generated by PCR and tested in an *in vitro* site-specific recombination assay. In this way all possible single base pair changes within the 7 bp overlap region of the core *att* site were generated as well as five additional changes outside the 7 bp overlap but within the 15 bp core *att* site. Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates.

10

### **Methods**

To examine both the efficiency and specificity of recombination of mutant *attL* and *attR* sites, a simple *in vitro* site-specific recombination assay was developed. Since the core regions of *attL* and *attR* lie near the ends of these sites, it was possible to incorporate the desired nucleotide base changes within PCR primers and generate a series of PCR products containing mutant *attL* and *attR* sites. PCR products containing *attL* and *attR* sites were used as substrates in an *in vitro* reaction with GATEWAY™ LR Clonase™ Enzyme Mix (Life Technologies, Inc.; Rockville, MD). Recombination between a 1.5 kb *attL* PCR product and a 1.0 kb *attR* PCR product resulted in a 2.5 kb recombinant molecule that was monitored using agarose gel electrophoresis and ethidium bromide staining.

Plasmid templates pEYC1301 (Figure 84) and pEYC1313 (Figure 85), each containing a single wild type *attL* or *attR* site, respectively, were used for the generation of recombination substrates. The following list shows primers that were used in PCR reactions to generate the *attL* PCR products that were used as substrates in L x R Clonase reactions (capital letters represent changes from the wild-type sequence, and the underline represents the 7 bp overlap region within the 15 bp core *att* site; a similar set of PCR primers was used to prepare the *attR* PCR products containing matching mutations):

25

30

-151-

GATEWAY™ sites (note: attL2 sequence in GATEWAY™ plasmids begins "acccca" while the attL2 site in this example begins "agcct" to reflect wild-type attL outside the core region.):

5

attL1: gggg agcct gctttttGtacAaa gttggcatta taaaaa-  
          agca ttgc

10

attL2: gggg agcct gctttCttGtacAaa gttggcatta taaaaa-  
          agca ttgc

Wild-type:

attL0: gggg agcct gctttttataactaa gttggcatta taaaaa-  
          agca ttgc

15

Single base changes from wild-type:

attLT1A: gggg agcct gctttAttataactaa gttggcatta taaaaa-  
          agca ttgc

20

attLT1C: gggg agcct gctttCttataactaa gttggcatta taaaaa-  
          agca ttgc

attLT1G: gggg agcct gctttGttataactaa gttggcatta taaaaa-  
          agca ttgc

25

attLT2A: gggg agcct gctttAtataactaa gttggcatta taaaaa-  
          agca ttgc

30

attLT2C: gggg agcct gctttCtataactaa gttggcatta taaaaa-  
          agca ttgc

attLT2G: gggg agcct gctttGtataactaa gttggcatta taaaaa-  
          aagca ttgc

35

-152-

attLT3A: gggg agcct gcttttAataactaa gttggcatta taaaa-  
aagca ttgc

5 attLT3C: gggg agcct gcttttCataactaa gttggcatta taaaa-  
aagca ttgc

10 attLT3G: gggg agcct gcttttGataactaa gttggcatta taaaa-  
aagca ttgc

15 attLA4C: gggg agcct gcttttCtactaa gttggcatta taaaa-  
aagca ttgc

20 attLA4G: gggg agcct gcttttGtactaa gttggcatta taaaa-  
aagca ttgc

25 attLA4T: gggg agcct gcttttTtactaa gttggcatta taaaa-  
aagca ttgc

attLT5A: gggg agcct gctttttaAactaa gttggcatta taaaa-  
aagca ttgc

30 attLT5C: gggg agcct gctttttaCactaa gttggcatta taaaa-  
aagca ttgc

35 attLT5G: gggg agcct gctttttaGactaa gttggcatta taaaa-  
aagca ttgc

attLA6C: gggg agcct gctttttatCctaa gttggcatta taaaa-  
aagca ttgc

-153-

attLA6G: gggg agcct gcttttatGctaa gttggcatta taaaa-  
aagca ttgc

5 attLA6T: gggg agcct gcttttatTctaa gttggcatta taaaa-  
aagca ttgc

10 attLC7A: gggg agcct gcttttataAataa gttggcatta taaaa-  
aagca ttgc

15 attLC7G: gggg agcct gcttttataGtaa gttggcatta taaaa-  
aagca ttgc

attLC7T: gggg agcct gcttttataTtaa gttggcatta taaaa-  
aagca ttgc

Single base changes outside of the 7 bp overlap:

20 attL8: gggg agcct Actttttataactaa gttggcatta taaaa-  
aagca ttgc

25 attL9: gggg agcct gcCttttataactaa gttggcatta taaaaaa-  
agca ttgc

attL10: gggg agcct gcttCtttataactaa gttggcatta taaaaaa-  
agca ttgc

30 attL14: gggg agcct gctttttatacCaa gttggcatta taaaaaa-  
agca ttgc

35 attL15: gggg agcct gctttttataactaG gttggcatta taaaaaa-  
agca ttgc

Note: additional vectors wherein the first nine bases are gggg agcca (*i.e.*, substituting an adenine for the thymine in the position immediately preceding the 15-bp core region), which may or may not contain the single base pair substitutions (or deletions) outlined above, can also be used in these experiments.

5

Recombination reactions of *attL*- and *attR*-containing PCR products was performed as follows:

8 µl of H<sub>2</sub>O  
10 2 µl of *attL* PCR product (100 ng)  
2 µl of *attR* PCR product (100 ng)  
4 µl of 5x buffer  
4 µl of GATEWAY™ LR Clonase™ Enzyme Mix  
20 µl total volume

15

Clonase reactions were incubated at 25°C for 2 hours.

2 µl of 10X Clonase stop solution (proteinase K, 2 mg/ml) were added to stop the reaction.

10 µl were run on a 1 % agarose gel.

20

### ***Results***

25

Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates. Changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination. These mutant *att* sites each recombined as well as the wild-type, but only with their cognate partner mutant; they did not recombine detectably with any other *att* site mutant. In contrast, changes in the last four positions (TTTATAC) only partially altered specificity; these mutants recombined with their cognate mutant as well as wild-type *att* sites and recombined partially with all other mutant *att* sites except for those having mutations in the first three positions of the 7 bp

30

-155-

overlap. Changes outside of the 7 bp overlap were found not to affect specificity of recombination, but some did influence the efficiency of recombination.

Based on these results, the following rules for *att* site specificity were determined:

5

- Only changes within the 7 bp overlap affect specificity.
- Changes within the first 3 positions strongly affect specificity.
- Changes within the last 4 positions weakly affect specificity.

Mutations that affected the overall efficiency of the recombination reaction were also assessed by this method. In these experiments, a slightly increased (less than 2-fold) recombination efficiency with *att*L T1A and *att*LC7T substrates was observed when these substrates were reacted with their cognate *att*R partners. Also observed were mutations that decreased recombination efficiency (approximately 2-3 fold), including *att*LA6G, *att*L14 and *att*L15. These mutations presumably reflect changes that affect Int protein binding at the core *att* site.

15 The results of these experiments demonstrate that changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination (*i.e.*, *att* sequences with one or more mutations in the first three thymidines would only recombine with their cognate partners and would not cross-react with any other *att* site mutation). In contrast, mutations in the last four positions (TTTATAC) only partially altered specificity (*i.e.*, *att* sequences with one or more mutations in the last four base positions would cross-react partially with the wild-type *att* site and all other mutant *att* sites, except for those having mutations in one or more of the first three positions of the 7 bp overlap). Mutations outside of the 7 bp overlap were not found to affect specificity of recombination, but some were found to influence (*i.e.*, to cause a decrease in) the efficiency of recombination.

25

***Example 22: Discovery of Att Site Mutations That Increase the Cloning Efficiency of GATEWAY™ Cloning Reactions***

30 In experiments designed to understand the determinants of *att* site specificity, point mutations in the core region of *att*L were made. Nucleic acid molecules containing these mutated *att*L sequences were then reacted in an LR

reaction with nucleic acid molecules containing the cognate *attR* site (*i.e.*, an *attR* site containing a mutation corresponding to that in the *attL* site), and recombinational efficiency was determined as described above. Several mutations located in the core region of the *att* site were noted that either slightly increased (less than 2-fold) or decreased (between 2-4-fold) the efficiency of the recombination reaction (Table 3).

*Table 3. Effects of attL mutations on Recombination Reactions.*

Site	Sequence	Effect on Recombination
attL0	agcctgcttttataactaagggtggcatta	
attL5	agcctgctttAtataactaagggtggcatta	slightly increased
attL6	agcctgcttttataTtaagggtggcatta	slightly increased
attL13	agcctgcttttatGctaagggtggcatta	decreased
attL14	agcctgcttttatacCaagggtggcatta	decreased
attL15	agcctgcttttataactaGgtggcatta	decreased
consensus	CAACTTnnTnnnAnnAAGTTG	

It was also noted that these mutations presumably reflected changes that either increased or decreased, respectively, the relative affinity of the integrase protein for binding the core *att* site. A consensus sequence for an integrase core-binding site (CAACTTNNT) has been inferred in the literature but not directly tested (see, *e.g.*, Ross and Landy, *Cell* 33:261-272 (1983)). This consensus core integrase-binding sequence was established by comparing the sequences of each of the four core *att* sites found in *attP* and *attB* as well as the sequences of five non-*att* sites that resemble the core sequence and to which integrase has been shown to bind *in vitro*. These experiments suggest that many more *att* site mutations might be identified which increase the binding of integrase to the core *att* site and thus increase the efficiency of GATEWAY™ cloning reactions.

**Example 23: Effects of Core Region Mutations on Recombination Efficiency**

To directly compare the cloning efficiency of mutations in the att site core region, single base changes were made in the attB2 site of an attB1-TET-attB2 PCR product. Nucleic acid molecules containing these mutated attB2 sequences were then reacted in a BP reaction with nucleic acid molecules containing non-cognate attP sites (*i.e.*, wildtype attP2), and recombinational efficiency was determined as described above. The cloning efficiency of these mutant attB2 containing PCR products compared to standard attB1-TET-attB2 PCR product are shown in Table 4.

5

10

**Table 4. Efficiency of Recombination With Mutated attB2 Sites.**

	<u>Site</u>	<u>Sequence</u>	<u>Mutation</u>	<u>Cloning Efficiency</u>
15	attB0	tcaagttagtataaaaaaggcaggct		
	attB1	ggggacaagttgtacaaaaaggcaggct		
	attB2	ggggaccacttgtacaagaaagctgggt		100%
20	attB2.1	ggggaAactttgtacaagaaagctgggt	C→A	40%
	attB2.2	ggggacAactttgtacaagaaagctgggt	C→A	131%
	attB2.3	ggggaccCctttgtacaagaaagctgggt	A→C	4%
	attB2.4	ggggaccaAtttgtacaagaaagctgggt	C→A	11%
	attB2.5	ggggaccacGttgtacaagaaagctgggt	T→G	4%
	attB2.6	ggggaccactGtgtacaagaaagctgggt	T→G	6%
25	attB2.7	ggggaccacttGgtacaagaaagctgggt	T→G	1%
	attB2.8	ggggaccacttTtacaagaaagctgggt	G→T	0.5%

As noted above, a single base change in the attB2.2 site increased the cloning efficiency of the attB1-TET-attB2.2 PCR product to 131% compared to the attB1-TET-attB2 PCR product. Interestingly, this mutation changes the integrase core binding site of attB2 to a sequence that matches more closely the proposed consensus sequence.

30

Additional experiments were performed to directly compare the cloning efficiency of an attB1-TET-attB2 PCR product with a PCR product that contained attB sites containing the proposed consensus sequence (see Example 22) of an integrase core binding site. The following attB sites were used to amplify attB-TET PCR products:

attB1 ggggacaagttgtacaaaaagcaggct  
 attB1.6 ggggacaactttgtacaaaaaagTTggct  
 attB2 ggggaccacttgtacaaqaaagctgggt  
 attB2.10 ggggacaacttgtacaaqaaagTtgggt

BP reactions were carried out between 300 ng (100 fmoles) of pDONR201 (Figure 49A) with 80 ng (80 fmoles) of attB-TET PCR product in a 20 µl volume with incubation for 1.5 hrs at 25°C, creating pENTR201-TET Entry clones. A comparison of the cloning efficiencies of the above-noted attB sites in BP reactions is shown in Table 5.

*Table 5. Cloning efficiency of BP Reactions.*

PCR product	CFU/ml	Fold Increase
B1-tet-B2	7,500	
B1.6-tet-B2	12,000	1.6 x
B1-tet-B2.10	20,900	2.8 x
B1.6-tet-B2.10	30,100	4.0 x

These results demonstrate that attB PCR products containing sequences that perfectly match the proposed consensus sequence for integrase core binding sites can produce Entry clones with four-fold higher efficiency than standard Gateway attB1 and attB2 PCR products.

The entry clones produced above were then transferred to pDEST20 (Figure 40A) via LR reactions (300 ng (64 fmoles) pDEST20 mixed with 50 ng (77 fmoles) of the respective pENTR201-TET Entry clone in 20 µl volume; incubated for 1 hr incubation at 25°C). The efficiencies of cloning for these reactions are compared in Table 6.

-159-

Table 6. Cloning Efficiency of LR Reactions.

pENTR201-TET x pDEST20	CFU/ml	Fold Increase
L1-tet-L2	5,800	
L1.6-tet-L2	8,000	1.4
L1-tet-L2.10	10,000	1.7
L1.6-tet-L2.10	9,300	1.6

5

10

15

These results demonstrate that the mutations introduced into attB1.6 and attB2.10 that transfer with the gene into entry clones slightly increase the efficiency of LR reactions. Thus, the present invention encompasses not only mutations in *attB* sites that increase recombination efficiency, but also to the corresponding mutations that result in the *attL* sites created by the BP reaction.

15

To examine the increased cloning efficiency of the attB1.6-TET-attB2.10 PCR product over a range of PCR product amounts, experiments analogous to those described above were performed in which the amount of attB PCR product was titrated into the reaction mixture. The results are shown in Table 7.

20

25

Table 7. Titration of attB PCR products.

Amount of attB PCR product (ng)	PCR product	CFU/ml	Fold Increase
20	attB1-TET-attB2	3,500	6.1
	attB1.6-TET-attB2.10	21,500	
50	attB1-TET-attB2	9,800	5.0
	attB1.6-TET-attB2.10	49,000	
100	attB1-TET-attB2	18,800	2.8
	attB1.6-TET-attB2.10	53,000	
200	attB1-TET-attB2	19,000	2.5
	attB1.6-TET-attB2.10	48,000	

These results demonstrate that as much as a six-fold increase in cloning efficiency is achieved with the attB1.6-TET-attB2.10 PCR product as compared to the standard attB1-TET-attB2 PCR product at the 20 ng amount.

30

**Example 24: Determination of attB Sequence Requirements for Optimum Recombination Efficiency**

To examine the sequence requirements for attB and to determine which attB sites would clone with the highest efficiency from populations of degenerate attB sites, a series of experiments was performed. Degenerate PCR primers were designed which contained five bases of degeneracy in the B-arm of the attB site. These degenerate sequences would thus transfer with the gene into Entry clone in BP reactions and subsequently be transferred with the gene into expression clones in LR reactions. The populations of degenerate attB and attL sites could thus be cycled from attB to attL back and forth for any number of cycles. By altering the reaction conditions at each transfer step (for example by decreasing the reaction time and/or decreasing the concentration of DNA) the reaction can be made increasingly more stringent at each cycle and thus enrich for populations of attB and attL sites that react more efficiently.

The following degenerate PCR primers were used to amplify a 500 bp fragment from pUC18 which contained the lacZ alpha fragment (only the attB portion of each primer is shown):

20 attB1 GGGG ACAAGTTGTACAAA AAAGC AGGCT  
attB1n16-20 GGGG ACAAGTTGTACAAA nnnnn AGGCT  
attB1n21-25 GGGG ACAAGTTGTACAAA AAAGC nnnnn

25 attB2 GGGG ACCACTTGTACAAG AAAGC TGGGT  
attB2n16-20 GGGG ACCACTTGTACAAG nnnnn TGGGT  
attB2n21-25 GGGG ACCACTTGTACAAG AAAGC nnnnn

30 The starting population size of degenerate att sites is  $4^5$  or 1024 molecules. Four different populations were transferred through two BP reactions and two LR reactions. Following transformation of each reaction, the population of transformants was amplified by growth in liquid media containing the appropriate selection antibiotic. DNA was prepared from the population of clones by alkaline

-161-

lysis miniprep and used in the next reaction. The results of the BP and LR cloning reactions are shown below.

BP-1, overnight reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	78,500	100 %
attB1n16-20-LacZa-attB2	1,140	1.5 %
attB1n21-25-LacZa-attB2	11,100	14 %
attB1-LacZa-attB2n16-20	710	0.9 %
attB1-LacZa-attB2n21-25	16,600	21 %

LR-1, pENTR201-LacZa x pDEST20/*EcoRI*, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	20,000	100 %
attL1n16-20-LacZa-attL2	2,125	11 %
attL1n21-25-LacZa-attL2	2,920	15 %
attL1-LacZa-attL2n16-20	3,190	16 %
attL1-LacZa-attL2n21-25	1,405	7 %

BP-2, pEXP20-LacZa/*ScaI* x pDONR 201, 1hr reactions.

	cfu/ml	percent of control
attB1-LacZa-attB2	48,600	100 %
attB1n16-20-LacZa-attB2	22,800	47 %
attB1n21-25-LacZa-attB2	31,500	65 %
attB1-LacZa-attB2n16-20	42,400	87 %
attB1-LacZa-attB2n21-25	34,500	71 %

LR-2, pENTR201-LacZa x pDEST6/*NcoI*, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	23,000	100 %
attL1n16-20-LacZa-attL2	49,000	213 %
attL1n21-25-LacZa-attL2	18,000	80 %
attL1-LacZa-attL2n16-20	37,000	160 %
attL1-LacZa-attL2n21-25	57,000	250 %

These results demonstrate that at each successive transfer, the cloning efficiency of the entire population of att sites increases, and that there is a great deal of flexibility in the definition of an *attB* site. Specific clones may be isolated from the above reactions, tested individually for recombination efficiency, and

sequenced. Such new specificities may then be compared to known examples to guide the design of new sequences with new recombination specificities. In addition, based on the enrichment and screening protocols described herein, one of ordinary skill can easily identify and use sequences in other recombination sites, e.g., other *att* sites, *lox*, FRT, etc., that result in increased specificity in the recombination reactions using nucleic acid molecules containing such sequences.

5

***Example 25: Design of att Site PCR Adapter-Primers***

10

Additional studies were performed to design gene-specific primers with 12bp of attB1 and attB2 at their 5'-ends. The optimal primer design for *att*-containing primers is the same as for any PCR primers: the gene-specific portion of the primers should ideally have a Tm of > 50°C at 50 mM salt (calculation of Tm is based on the formula  $59.9 + 41(\%GC) - 675/n$ ).

15

Primers:

12bp attB1: AA AAA GCA GGC TNN - forward gene-specific primer

20

12bp attB2: A GAA AGC TGG GTN - reverse gene-specific primer

attB1 adapter primer: GGGGACAAGTTGTACAAAAAAGCAGGCT

25

Protocol:

(1) Mix 200 ng of cDNA library or 1 ng of plasmid clone DNA (alternatively, genomic DNA or RNA could be used) with 10 pmoles of gene specific primers in a 50 µl PCR reaction, using one or more polypeptides having DNA polymerase activity such as those described herein. (The addition of greater than 10 pmoles of gene-specific primers can decrease the yield of attB PCR product. In addition, if RNA is used, a standard reverse transcriptase-PCR (RT-

30

PCR) protocol should be followed; *see, e.g.*, Gerard, G.F., *et al.*, *FOCUS* 11:60 (1989); Myers, T.W., and Gelfand, D.H., *Biochem.* 30:7661 (1991); Freeman, W.N., *et al.*, *BioTechniques* 20:782 (1996); and U.S. Application No. 09/064,057, filed April 22, 1998, the disclosures of all of which are incorporated herein by reference.)

5

1<sup>st</sup> PCR profile:

- (a) 95°C for 3 minutes
- (b) 10 cycles of:
  - (i) 94°C for 15 seconds
  - (ii) 50°C\* for 30 seconds
  - (iii) 68°C for 1 minute/kb of target amplicon
- (c) 68°C for 5 minutes
- (d) 10°C hold

10

15

\*The optimal annealing temperature is determined by the calculated Tm of the gene-specific part of the primer.

20 (2) Transfer 10 µl to a 40 µl PCR reaction mix containing 35 pmoles each of the attB1 and attB2 adapter primers.

25

2<sup>nd</sup> PCR profile:

- (a) 95°C for 1 minute
- (b) 5 cycles of:
  - (i) 94°C for 15 seconds
  - (ii) 45°C\* for 30 seconds
  - (iii) 68°C for 1 minute/kb of target amplicon
- (c) 15-20 cycles\*\* of:
  - (i) 94°C for 15 seconds
  - (ii) 55°C\* for 30 seconds

30

-164-

- (iii) 68°C for 1 minute/kb of target amplicon
- (d) 68°C for 5 minutes
- (e) 10°C hold

5 \*The optimal annealing temperature is determined by the calculated Tm of the gene-specific part of the primer.

\*\*15 cycles is sufficient for low complexity targets.

Notes:

10 1. It is useful to perform a no-adapter primer control to assess the yield of attB PCR product produced.

2. Linearized template usually results in slightly greater yield of PCR product.

15

*Example 26: One-Tube Recombinational Cloning Using the GATEWAY™ Cloning System*

20 To provide for easier and more rapid cloning using the GATEWAY™ cloning system, we have designed a protocol whereby the BP and LR reactions may be performed in a single tube (a "one-tube" protocol). The following is an example of such a one-tube protocol; in this example, an aliquot of the BP reaction is taken before adding the LR components, but the BP and LR reactions may be performed in a one-tube protocol without first taking the BP aliquot:

25

<u>Reaction Component</u>	<u>Volume</u>
attB DNA (100-200 ng/25 µl reaction)	1-12.5 µl
attP DNA (pDONR201) 150 ng/µl	2.5 µl
5X BP Reaction Buffer	5.0 µl
Tris-EDTA	(to 20 µl)
<u>BP Clonase</u>	<u>5.0 µl</u>
Total vol.	25 µl

-165-

After the above components were mixed in a single tube, the reaction mixtures were incubated for 4 hours at 25°C. A 5 µl aliquot of reaction mixture was removed, and 0.5 µl of 10X stop solution was added to this reaction mixture and incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the BP reaction per 100 µl of cells; this transformation yielded colonies of Entry Clones for isolation of individual Entry Clones and for quantitation of the BP Reaction efficiency.

To the remaining 20 µl of BP reaction mixture, the following components of the LR reaction were added:

<u>Reaction Component</u>	<u>Final Concentration</u>	<u>Volume Added</u>
NaCl	0.75 M	1 µl
Destination Vector	150 ng/µl	3 µl
<u>LR Clonase</u>		<u>6 µl</u>
Total vol.		30 µl

After the above components were mixed in a single tube, the reaction mixtures were incubated for 2 hours at 25°C. 3 µl of 10X stop solution was added, and the mixture was incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the reaction mixture per 100 µl of cells

Notes:

1. If desired, the Destination Vector can be added to the initial BP reaction.
2. The reactions can be scaled down by 2x, if desired.
3. Shorter incubation times for the BP and/or LR reactions can be used (scaled to the desired cloning efficiencies of the reaction); but a lower number of colonies will typically result.
4. To increase the number of colonies obtained by several fold, incubate the BP reaction for 6-20 hours and increase the LR reaction to 3 hours. Electroporation also works well with 1-2 ul of the PK-treated reaction mixture.

5. PCR products greater than about 5 kb may show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using a one-tube reaction with longer incubation times (e.g., 6-18 hours) for both the BP and LR steps.

5

***Example 27: Relaxation of Destination Vectors During the LR Reaction***

To further optimize the LR Reaction, the composition of the LR Reaction buffer was modified from that described above and this modified buffer was used  
10 in a protocol to examine the impact of enzymatic relaxation of Destination Vectors during the LR Reaction.

LR Reactions were set up as usual (see, e.g., Example 6), except that 5X BP Reaction Buffer (see Example 5) was used for the LR Reaction. To accomplish Destination Vector relaxation during the LR Reaction,  
15 Topoisomerase I (Life Technologies, Inc., Rockville, MD; Catalogue No. 38042-016) was added to the reaction mixture at a final concentration of ~15U per  $\mu$ g of total DNA in the reaction (for example, for reaction mixtures with a total of 400ng DNA in the 20  $\mu$ l LR Reaction, ~6units of Topoisomerase I was added). Reaction mixtures were set up as follows:

20

<u>Reaction Component</u>	<u>Volume</u>
ddH <sub>2</sub> O	6.5 $\mu$ l
4X BP Reaction Buffer	5 $\mu$ l
100ng single chain/linear pENTR CAT, 50 ng/ $\mu$ l	2 $\mu$ l
25 300ng single chain/linear pDEST6, 150ng/ $\mu$ l	2 $\mu$ l
Topoisomerase I, 15 U/ml	0.5 $\mu$ l
LR Clonase	4 $\mu$ l

30 Reaction mixtures were incubated at 25°C for 1hour, and 2  $\mu$ l of 2  $\mu$ g/ $\mu$ l Proteinase K was then added and mixtures incubated for 10 minutes at 37°C to stop the LR Reaction. Competent cells were then transformed as described in the preceding examples. The results of these studies demonstrated that relaxation of

-167-

substrates in the LR reaction using Topoisomerase I resulted in a 2- to 10-fold increase in colony output compared to those LR reactions performed without including Topoisomerase I.

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

167.1

Applicant's or agent's file reference number	0942.~8PC03	International application No. t <sup>1</sup> PCT/US 00/05432
--	-------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL**  
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 52, line 31.

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

## Name of depositary institution

Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30099
--------------------------------------	----------------------------------

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet 

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

## For receiving Office use only

This sheet was received with the international application

Authorized officer *[Signature]*  
USPTO - 1709 Tenth Street, Suite 1000  
Washington, DC 20036-3220 (PA)

## For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

167.2

Applicant's or agent's file reference number	International application No. t <sub>1</sub> PCT/US 00/05432
0942.468PC03	

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL  
(PCT Rule 13bis)**

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution  
Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30100
--------------------------------------	----------------------------------

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet 

Escherichia coli DB3.1(pENTR-1A)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

 This sheet was received with the international application

Authorized officer

Barbara Frisch  
U.S. Patent and Trademark Office  
1100 L Street, NW  
Washington, DC 20591-0000

For International Bureau use only

 This sheet was received by the International Bureau on:

Authorized officer

167 3	
Applicant's or agent's file reference number	International application No. tb PCT/US00/05432
0942.468PC03	

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL  
(PCT Rule 13bis)**

REC'D 1. 11. 1990

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 16.

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

## Name of depositary institution

Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

## Date of deposit

February 27, 1999

## Accession Number

NRRL B-30101

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet 

Escherichia coli DB3.1(pENTR-2B)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer <i>[Signature]</i> J. S. Garcia, P.T.C. IT Operations Manager WIPO, P.O. Box 3-3230 (P)	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

167.4

Applicant's or agent's file reference number	0942.468PC03	International Application No. t <sub>1</sub> PCT/US00/05432
--	--------------	--

REC'D 17 APR 2000

MATERIAL

WIPO

PCT

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL  
(PCT Rule 13bis)**

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet 

Name of depositary institution  
Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

Date of deposit  
February 27, 1999

Accession Number  
NRRL B-30102

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)

This information is continued on an additional sheet 

Escherichia coli DB3.1(pENTR-3C)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

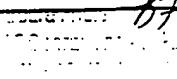
The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

 This sheet was received with the international application This sheet was received by the International Bureau on:

Authorized officer

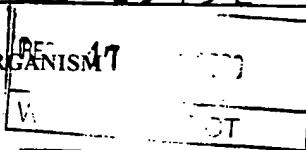

  
 U.S. Patent and Trademark Office  
 PCT/US00/05432

Authorized officer

167.5

Applicant's or agent's file reference number	0942.468PC03	International application No. tb. PCT/US 00/05432
--	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL  
(PCT Rule 13bis)**



A. The indications made below relate to the microorganism referred to in the description on page <u>8</u> .	REC'D 17 APR 2000
---	-------------------

<b>B. IDENTIFICATION OF DEPOSIT</b>	Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	

Address of depositary institution ( <i>including postal code and country</i> )  1815 N. University Street Peoria, Illinois 61604 United States of America	
---	--

Date of deposit February 27, 1999	Accession Number NRRL B-30103
--------------------------------------	----------------------------------

<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )	This information is continued on an additional sheet <input type="checkbox"/>
Escherichia coli DB3.1(pEZC15101)	

<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )
--

<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the international Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )
---

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer  FEB 27 1999 WIPO PCT/US00/05432 (E)	Authorized officer

167.6

Applicant's or agent's file reference number 0942.468PC03	International application No. t. PCT/US 10/05432
--	---

REC'D 17

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL *WPO***  
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution ( <i>including postal code and country</i> )  1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30104
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )	
This information is continued on an additional sheet <input type="checkbox"/>  Escherichia coli DB3.1(pEYC15102)	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )	
The indications listed below will be submitted to the international Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	<i>[Signature]</i>	Authorized officer	

167.7

Applicant's or agent's file reference number	0942.468PC03	International application No. <b>PCT/US 00/05432</b>
--	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL**  
(PCT Rule 13bis)

PCT/US 00/05432 ARR 100

V T

A. The indications made below relate to the microorganism referred to in the description on page 54, line 9.

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution  
Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30105
--------------------------------------	----------------------------------

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet 

Escherichia coli DB3.1(pEZA15103)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

 This sheet was received with the international application

 This sheet was received by the International Bureau on:

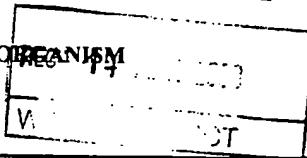
Authorized officer *Barbara Frisch* *BF*  
*International Bureau of the World Intellectual Property Organization*  
*36 Avenue de la Paix, 1211 Geneva 20, Switzerland*

Authorized officer

167.8

Applicant's or agent's file reference number	0942.408PC03	International application No. tl PCT/US 00/05432
--	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL**  
(PCT Rule 13bis)



A. The indications made below relate to the microorganism referred to in the description on page 51, line 20-21.

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution  
Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30108
--------------------------------------	----------------------------------

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)

This information is continued on an additional sheet

Escherichia coli DB10B(pCMV Sport6)

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

Barbara Fries

AS-102242  
PCT/US00/05432  
00/05432

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

## WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group of nucleotide sequences consisting of an attB1 nucleotide sequence as set forth in Figure 9, an attB2 nucleotide sequence as set forth in Figure 9, an attP1 nucleotide sequence as set forth in Figure 9, an attP2 nucleotide sequence as set forth in Figure 9, an attL1 nucleotide sequence as set forth in Figure 9, an attL2 nucleotide sequence as set forth in Figure 9, an attR1 nucleotide sequence as set forth in Figure 9, an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, and a mutant, fragment, or derivative thereof.
2. An isolated nucleic acid molecule comprising an attB1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
3. An isolated nucleic acid molecule comprising an attB2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
4. An isolated nucleic acid molecule comprising an attP1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
5. An isolated nucleic acid molecule comprising an attP2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
6. An isolated nucleic acid molecule comprising an attL1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

7. An isolated nucleic acid molecule comprising an attL2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

5 8. An isolated nucleic acid molecule comprising an attR1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

10 9. An isolated nucleic acid molecule comprising an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

15 10. The isolated nucleic acid molecule of claim 1, further comprising one or more functional or structural nucleotide sequences selected from the group consisting of one or more multiple cloning sites, one or more localization signals, one or more transcription termination sites, one or more transcriptional regulatory sequences, one or more translational signals, one or more origins of replication, one or more fusion partner peptide-encoding nucleic acid molecules, one or more protease cleavage sites, and one or more 5' polynucleotide extensions.

20 11. The nucleic acid molecule of claim 10, wherein said transcriptional regulatory sequence is a promoter, an enhancer, or a repressor.

25 12. The nucleic acid molecule of claim 10, wherein said fusion partner peptide-encoding nucleic acid molecule encodes glutathione S-transferase (GST), hexahistidine (His<sub>6</sub>), or thioredoxin (Trx).

13. The nucleic acid molecule of claim 10, wherein said 5' polynucleotide extension consists of from one to five nucleotide bases.

30 14. The nucleic acid molecule of claim 13, wherein said 5' polynucleotide extension consists of four or five guanine nucleotide bases.

-170-

15. A primer nucleic acid molecule suitable for amplifying a target nucleotide sequence, comprising the isolated nucleic acid molecule of claim 1 or a portion thereof linked to a target-specific nucleotide sequence useful in amplifying said target nucleotide sequence.

5

16. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB1 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

10

17. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB2 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

15

18. The primer nucleic acid molecule of claim 15, further comprising a 5' terminal extension of four or five guanine bases.

20

19. A vector comprising the isolated nucleic acid molecule of claim 1.

20. The vector of claim 19, wherein said vector is an Expression Vector.

25

21. A host cell comprising the isolated nucleic acid molecule of claim 1 or the vector of claim 19.

30

22. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

(a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said

templates and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and

5 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

10

23. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

(a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said templates and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and

(b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

20

25

30 24. A method of amplifying or synthesizing one or more nucleic acid molecules comprising:

(a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity

and one or more first primers comprising at least a portion of a recombination site and a template-specific sequence that is complementary to or capable of hybridizing to said template;

5 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one or both termini of said molecules;

10 (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and

15 (d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one or both termini of said molecules.

20 25. A polypeptide encoded by the isolated nucleic acid molecule of any one of claims 1-10.

25 26. An isolated nucleic acid molecule comprising one or more *att* recombination sites comprising at least one mutation in its core region that increases the specificity of interaction between said recombination site and a second *att* recombination site.

30 27. The isolated nucleic acid molecule of claim 26, wherein said mutation is at least one substitution mutation of at least one nucleotide in the seven basepair overlap region of said core region of said recombination site.

28. The isolated nucleic acid molecule of claim 26, wherein said nucleic acid molecule comprises the sequence NNNATAC, wherein "N" refers to any nucleotide with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

5

29. An isolated nucleic acid molecule comprising one or more mutated *att* recombination sites comprising at least one mutation in its core region that enhances the efficiency of recombination between a first nucleic acid molecule comprising said mutated *att* recombination site and a second nucleic acid molecule comprising a second recombination site that interacts with said mutated *att* recombination site.

10

30. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site is a mutated *attL* site comprising a core region having the nucleotide sequence caactnnntnnnnannaagttg, wherein "n" represents any nucleotide.

15

31. The isolated nucleic acid molecule of claim 30, wherein said mutated *attL* recombination site comprises a core region having a nucleotide sequence selected from agcctgcttattatactaaggtaatggcatta (*attL5*) and agcctgcttttatattaaggtaatggcatta (*attL6*).

20

32. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site comprises a core region having a nucleotide sequence selected from the group consisting of ggggacaacttgtacaaaaaagttggct (*attB*1.6), ggggacaacttgtacaagaaagctgggt (*attB*2.2), and ggggacaacttgtacaagaaagtgggt (*attB*2.10).

25

33. A vector selected from the group consisting of pENTR1A, pENTR2B, pENTR3C, pENTR4, pENTR5, pENTR6, pENTR7, pENTR8, pENTR9, pENTR10, pENTR11, pDEST1, pDEST2, pDEST3, pDEST4,

30

-174-

pDEST5, pDEST6, pDEST7, pDEST8, pDEST9, pDEST10, pDEST11, pDEST12.2 (also known as pDEST12), pDEST13, pDEST14, pDEST15, pDEST16, pDEST17, pDEST18, pDEST19, pDEST20, pDEST21, pDEST22, pDEST23, pDEST24, pDEST25, pDEST26, pDEST27, pDEST28, pDEST29, pDEST30, pDEST31, pDEST32, pDEST33, pDEST34, pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector), pDONR202, pDONR203 (also known as pEZ15812), pDONR204, pDONR205, pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector), pDONR207, pMAB58, pMAB62, pMAB85 and pMAB86.

5

10

34. A host cell comprising the vector of claim 33.

35. A polypeptide encoded by the vector of claim 33.

15

36. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the isolated nucleic acid molecule of any one of claims 1-10, 26 and 29.

20

37. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the primer of claim 15 or claim 18.

38. A kit for use in cloning a nucleic acid molecule, said kit comprising the vector of claim 19 or claim 33.

1/240

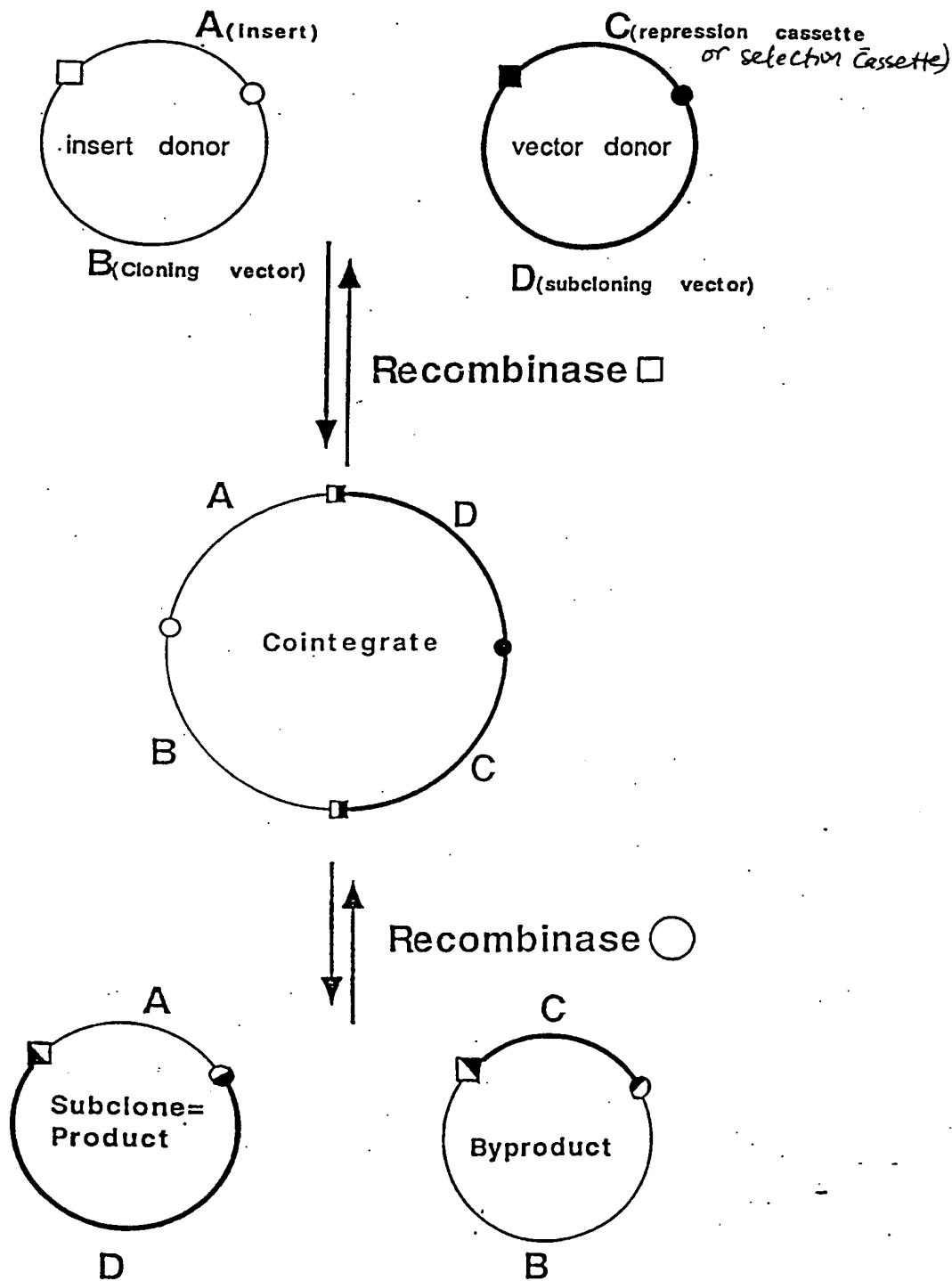


Figure 1

2/240

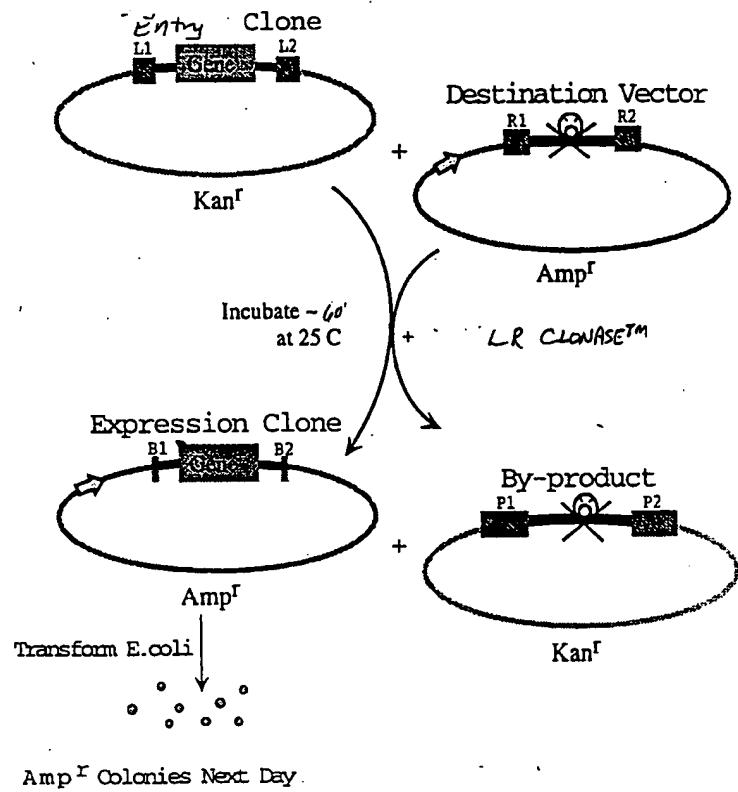


FIGURE 2

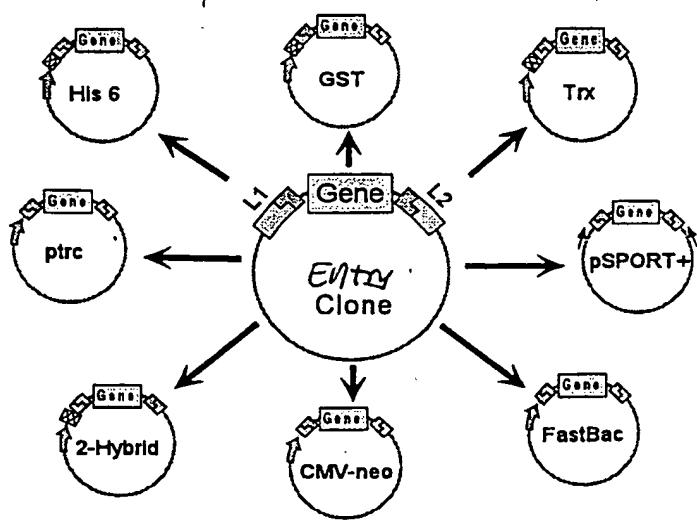


FIGURE 3

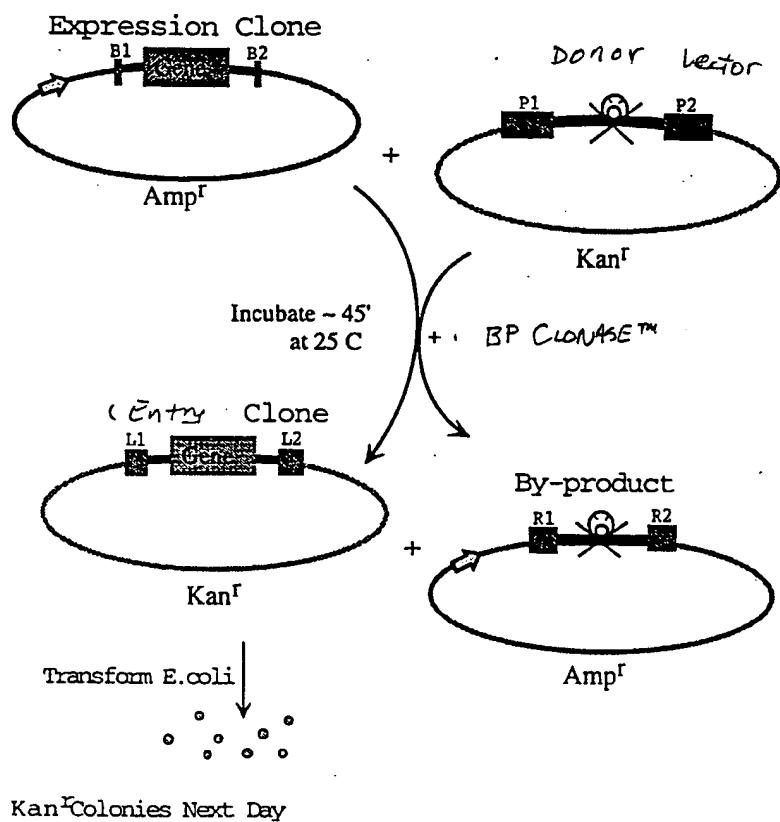


FIGURE 4

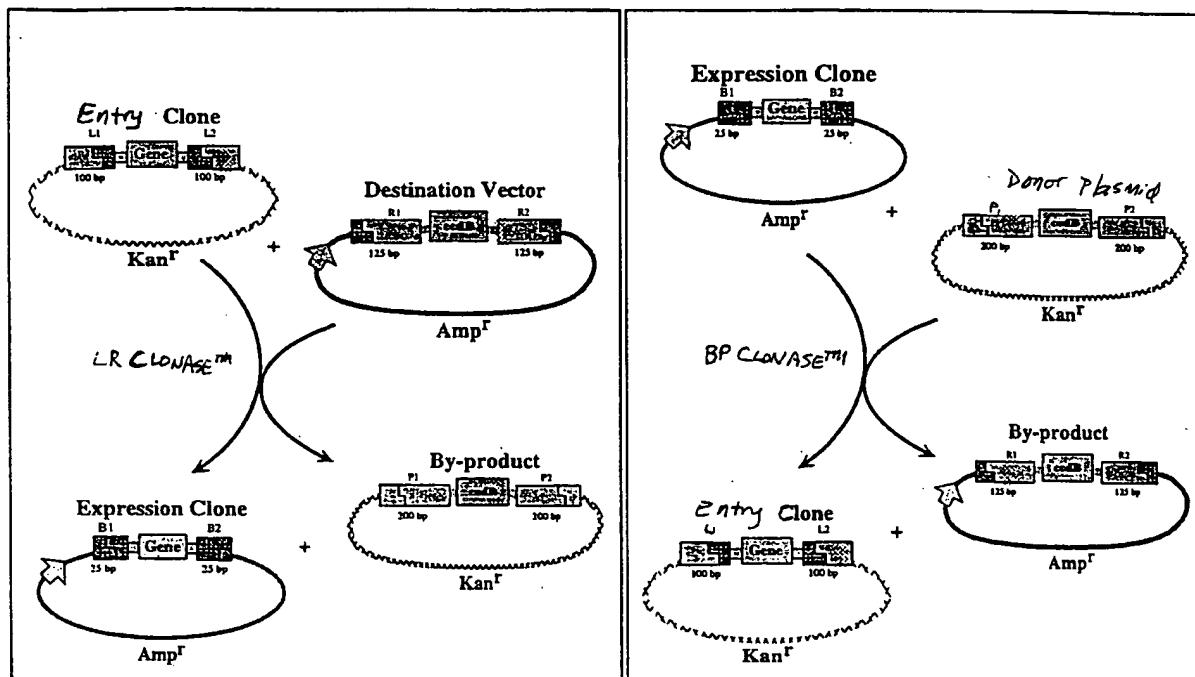
*A**B*

FIGURE 5

6/240

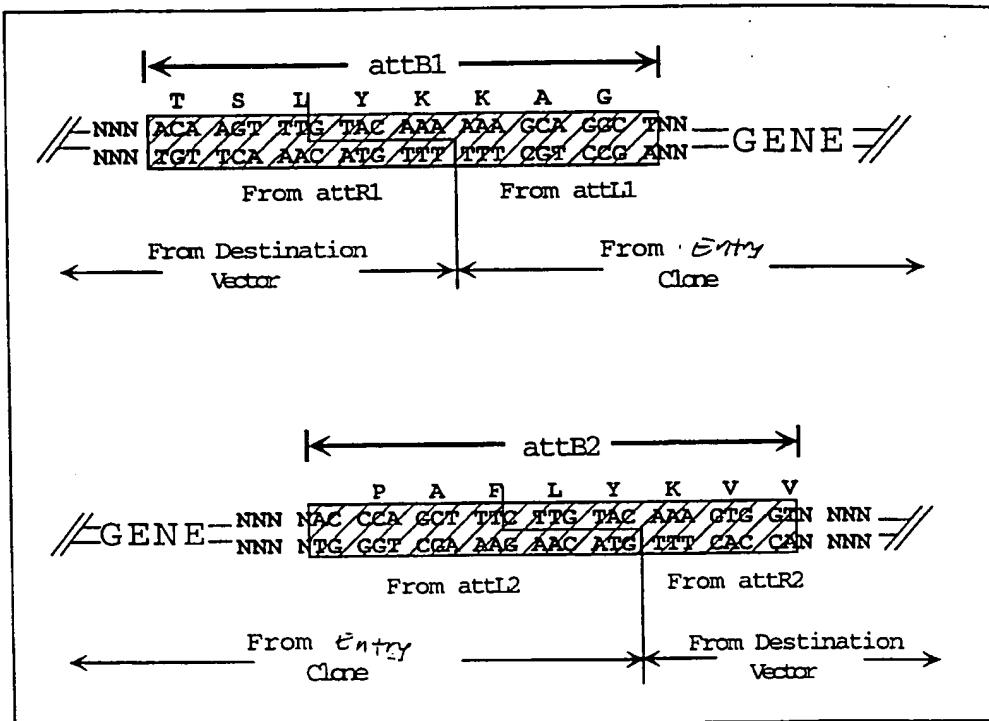


FIGURE 6

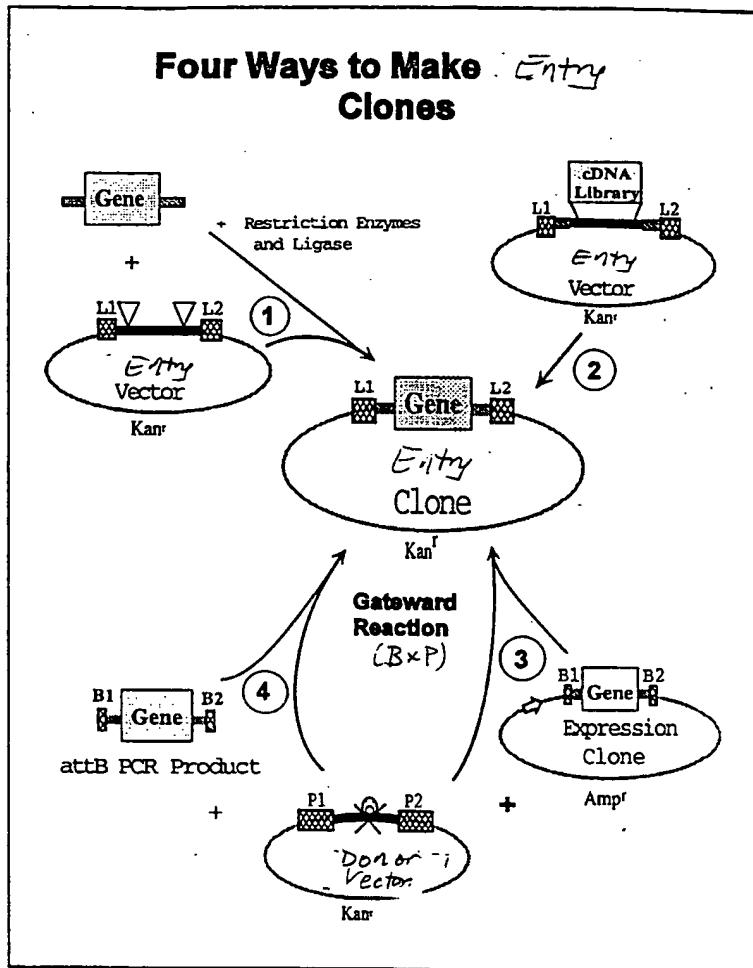


FIGURE 7

8/240

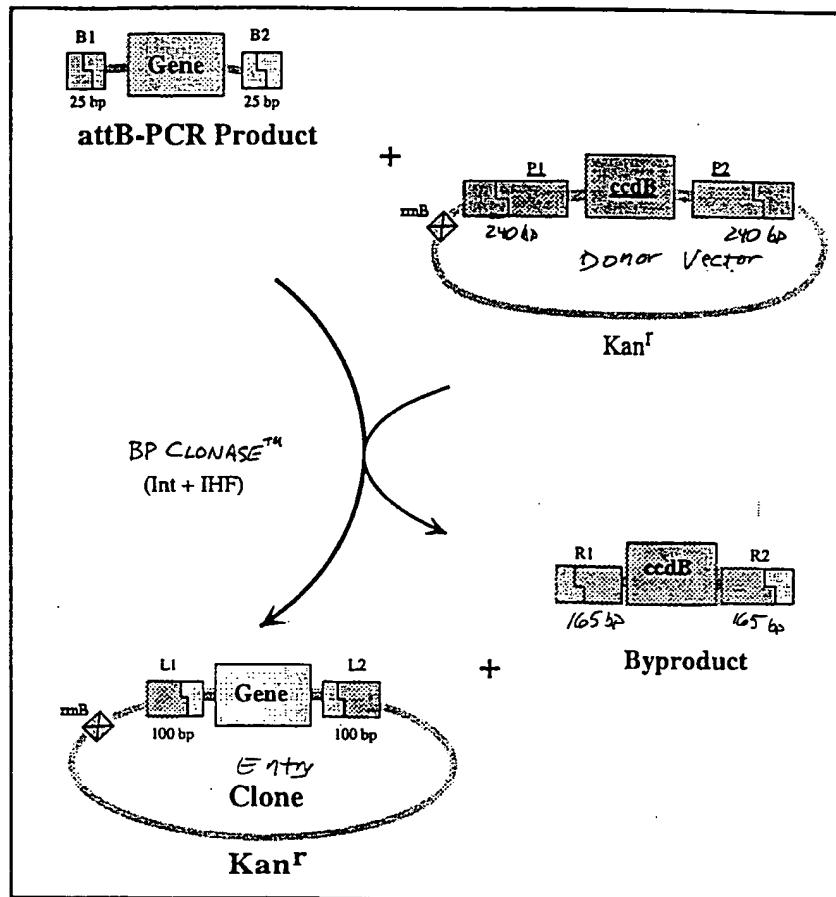


FIGURE 8

### Recombination Site Nucleotide Sequences

attB1: 5'-ACAAGTTGTACAAAAAAGCAGGCT-3'

attB2: 5'-ACCCAGCTTCTTGTACAAAGTGGT-3'

attP1: 5'-TACAGGTCACTAACCATCTAAGTAGTTGATTGACTGGATATG-TTGTTTACAGTATTATGTAGTCTGTTTATGCAAATCTAATTAA-TATATTGATATTATATCATTACGTTCTCGTTAGCTTTGTAC-AAAGTTGGCATTATAAAAAGCATTGCTCATCAATTGTTGCAACGAAC-GTCACTATCAGTCAAATAAAATCATTATTG-3'

attP2: 5'-CAAATAATGATTATTTGACTGATAGTGACCTGTTGCAACAAAT-TGATAAGCAATGCTTCTTATAATGCCAAGTTGACAAGAAAGCTGAAC-GAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTGCAT-AAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCACTATGA-ATCAACTACTTAGATGGTATTAGTGACCTGTA-3'

attR1: 5'-ACAAGTTGTACAAAAAAGCTAACGAGAAACGTAAAATGATATAAA-TATCAATATATTAAATTAGATTTGCATAAAAACAGACTACATAATAC-TGTTAAACACAACATATCCAGTCACTATG-3'

attR2: 5'-GCAGGTGACCATAGTGACTGGATATGTTGTTTACAGTATTAT-GTAGTCTGTTTATGCAAATCTAATTAAATATATTGATATTT-ATATCATTACGTTCTCGTTAGCTTTGTACAAAGTGGT-3'

attL1: 5'-CAAATAATGATTATTTGACTGATAGTGACCTGTTGCAAC-AAATTGATAAGCAATGCTTCTTATAATGCCAAGTTGACAAAAAA-GCAGGCT-3'

attL2: 5'-CAAATAATGATTATTTGACTGATAGTGACCTGTTGCAACAA-ATTGATAAGCAATGCTTCTTATAATGCCAAGTTGACAAAGAAAGCTGGGT-3'

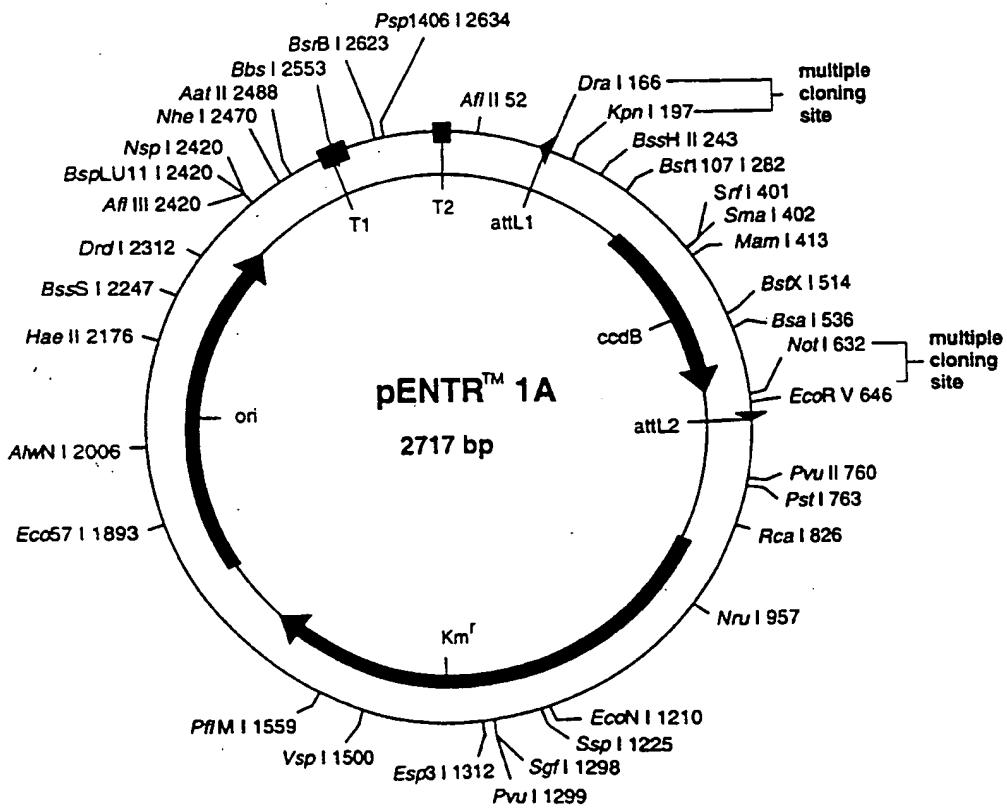
Figure 9

10/240

**Figure 10A: Cloning sites of the Entry Vector pENTR™ 1A (reading frame A)**

$\begin{matrix} \boxed{Dra\ I} & \boxed{Xmn\ I} & \boxed{Sal\ I} & \boxed{BamH\ I} & \boxed{Kpn\ I} & \boxed{EcoR\ I} \end{matrix}$   
 ACT TTG TAC AAA AAA GCA GGC TTT | AAA GGA ACC | AAT TCA GTC GAC TGG ATC CGG TAC | CGA ATT C  
 TGA AAC ATG TTT TTT CGT CCG AAA TTT CCT TGG TTA AGT CAG CTG ACC TAG | GCC ATG GCT TAA | G  
 thr leu tyr lys lys ala gly phe lys gly thr asn ser val asp trp ile arg tyr arg ile

$\begin{matrix} \boxed{EcoR\ I} & \boxed{Not\ I} & \boxed{Xho\ I} & \boxed{EcoR\ V} \end{matrix}$   
 --- [ccdB gene] --- G AAT TCG CGG CCG CAC TCG AGA TAT CTA GAC CCA GCT TTC TTG TAC AAA  
 C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT CGA AAG AAC ATG TTT



11/240

## pENTR1A 2717 bp

<u>Base Nos.</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAATTG TGACAAAAAG CAGGCTTTAA AGGAACCAAT  
 181 TCAGTCGACT GGATCCGGA CGGAATTGCG TTACTAAAAG CCAGATAACA GTATGCGTAT  
 241 TTGCGCGCTG ATTTTGCGG TATAAGAATA TATACTGATA TGTATACCG AAGTATGTCA  
 301 AAAAGAGGTG TGCTTCTAGA ATGCAGTTA ACGTTTACAC CTATAAAAGA GAGAGCCGTT  
 361 ATCGTCTGTT TGTGGATGTA CAGACTGATA TTATTGACAC GCCCAGGGCGA CGGATAGTGA  
 421 TCCCCCTGGC CAGTCACGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG  
 481 TGCATATCGG GGATGAAAGC TGACCATG TGACCCACCGA TATGGCCAGT GTGCCGGTCT  
 541 CCGTTATCGG GGAAGAACGT GCTGATCTCA GCCACCGCGA AAATGACATC AAAAACGCCA  
 601 TTAACCTGAT GTTCTGGGA ATATAGAATT CGCGGGCGCA CTCGAGATAT CTAGACCCAG  
 661 CTTCTTGTA CAAAGTTGGC ATTATAAGAA AGCATTTGCTT ATCAATTGTT TGCAACGAAC  
 721 AGGTCACTAT CAGTCAAAAT AAAATCATTA TTGCCCCATCC AGCTGCACT CTGGCCCGTG  
 781 TCTCAAATC TCTGATGTTA CATTGCAACAA GATAAAAATA TATCATCATG AACAAATAAA  
 841 CTGCTCTGCTT ACATAAACAG TAATACAAGG GGTGTTATGA GCCATATTCA ACGGGAAACG  
 901 TCGAGGCCGC GATTAAATTG CAAACATGGAT GCTGATTAT ATGGGTATAA ATGGGCTCGC  
 961 GATAATGTCG GGCAATCAGG TGCGACAATC TATCGTTGT ATGGGAAGCC CGATGCGCCA  
 1021 GAGTTGTTTC TGAAACATGG CAAAGGTAGC GTTGCCTAATG ATGTTACAGA TGAGATGGTC  
 1081 AGACTAAACT GGCTGACCGA ATTTATGCCT CTTCCGACCA TCAAGCATT TATCCGTACT  
 1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCCGGAA AAACAGCATT CCAGGTATT  
 1201 GAAGAAATATC CTGATTCAAG TGAAAATATT GTTGATGCGC TGGCAGTGTC CCTGCGCCGG  
 1261 TTGCATTGCA TTCTGTTTG TAATTGTCCT TTAAACAGCG ATCGCGTATT TCGTCTCGCT  
 1321 CAGGCAGCAAT CACGAATGAA TAACGGTTTG GTTGATGCGA GTGATTTGA TGACGAGCGT  
 1381 AATGGCTGGC CTGTTGAACA AGTCTGGAAA GAAATGCATA AACTTTGCC ATTCTCACCG  
 1441 GATTCAGTCG TCACTCATGG TGATTCTCA CTTGATAACC TTATTTTGA CGAGGGAAA  
 1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATACCA GGATCTTGC  
 1561 ATCCTATGGA ACTGCCTCGG TGAGTTTCT CCTTCATTAC AGAAACGGCT TTTCAAAAA  
 1621 TATGGTATTG ATAATCCTGA TATGAATAAA TTGCACTGTTT ATTTGATGCT CGATGAGTTT  
 1681 TTCTAATCAG AATTGGTTAA TTGGTTGTA CATTATTGAG ATGGGGCCCC GTTCCACTGA  
 1741 CGCTCAGACC CGCTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCCTTTT TCTGCGCGTA  
 1801 ATCTGCTGCT TGCAAAACAAA AAAACCACCG CTACCAAGCGG TGGTTGTTT GCCGGATCAA  
 1861 GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA GAGCGCAGAT ACCAAATACT  
 1921 GTTCTTCTAG TGTAGCCGA GTTACCGCAC CACTTCAAGA ACTCTGTAGC ACCGCCCTACA  
 1981 TACCTCGCTC TGCTAATCCT GTTACCGAGC GCTGCTGCCA GTGGCGATAA GTCGTGTCTT  
 2041 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTGGGG CTGAACGGGG  
 2101 GTTCTCGCA CACAGCCAG CTTGGAGCGA ACCACCTACA CGAACCTGAG ATACCTACAG  
 2161 CGTGAGCTAT GAGAAAGCGC CACGCCCTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA  
 2221 AGGGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT  
 2281 CTTTATAGTC CTGTCGGTT TGCCACCTC TGACTTGAGC GTCGATTGTTT GTGATGCTCG  
 2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACCGGG CCTTTTTACG GTTCCCTGGCC  
 2401 TTTTGCTGGC CTTTTGCCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC  
 2461 CGTATTACCG CTAGCATGGA TCTCGGGGAC GTCTAACTAC TAAGCGAGAG TAGGAACTG  
 2521 CCAGGCATCA AATAAAACGA AAGGCTCAGT CGGAAGACTG GGCCTTCGT TTTATCTGTT  
 2581 GTTTGTCGGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG  
 2641 TGAAGCAACG GCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGCATCAAA  
 2701 CTAAGCAGAA GGCCATC

FIGURE 10B

12/24/00

**Figure 11A: Cloning Sites of the Entry Vector pENTR2B (reading frame B)**

Int	attL1	EheI	XmnI	SalI	BamHI
-----	-------	------	------	------	-------

TTG TAC AAA AAA GCA GGC TGG CGC CGG AAC CAA TTC AGT CGA CTG GAT CCG  
 AAC ATG TTT TTT CGT CCG ACC GCG GCC TTG GTT AAG TCA GCT GAC CTA GGC  
 ↓  
 Leu Tyr Lys Lys Ala Gly Trp Arg Arg Asn Gln Phe Ser Arg Leu Asp Pro

KpnI	EcoRI	EcoRI	NotI	XbaI	EcoRV
------	-------	-------	------	------	-------

GTA DCG AAT TC- ccdB --G AAT TCG DGG CCG CAC TCG AGA TAT CTA GAC CCA  
 CAT GGC TTA AG C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT  
 ↓  
 Val Pro Asn Asn Ser Arg Pro His Ser Arg Tyr Leu Asp Pro

Int	attL2
-----	-------

GCT TTG TTG TAC AAA G  
 CGA AAG AAC ATG TTT C  
 ↓ ↓ ↓ ↓ ↓ ↓  
 Ala Phe Leu Tyr Lys

13/240

## pENTR2B 2718 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
322..627	ccdB
656..755	attL2
878..1687	KmR
1792..2365	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTG ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTGGCG CCGGAACCAA  
 181 TTCAGTCGAC TGGATCCGGT ACCGAATTG CTTACTAAAAA GCCAGATAAC AGTATGCGTA  
 241 TTTGCGCGCT GATTTTTGCG GTATAAGAAT ATATACTGAT ATGTATAACCC GAAGTATGTC  
 301 AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA CCTATATAAAAG AGAGAGCCGT  
 361 TATCGTCTGT TTGTGGATGT ACAGAGTGT ATTATTGACAA CGCCCGGGCG ACGGATGGTG  
 421 ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT CCCGTGAACCT TTACCCGGTG  
 481 GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG TGTGCCGGTC  
 541 TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG AAAATGACAT CAAAAACGCC  
 601 ATTAACCTGA TGTTCTGGGG AATATAGAAT TCGCGGCCGC ACTCGAGATA TCTAGACCCA  
 661 GCTTCTTGT ACAAAAGTTGG CATTATAAGA AAGCATTGCT TATCAATTG TTGCAACGAA  
 721 CAGGTCACTA TCAGTCAAAAA TAAAATCATT ATTGCCCAC CAGCTGCAGC TCTGGCCCGT  
 781 GTCTCAAAAT CTCTGATGTT ACATTGACA AGATAAAAAT ATATCATCAT GAACAATAAA  
 841 ACTGCTCTGCT TACATAAAACA GTAATACAAG GGGTGTATG AGCCATATTG AACGGGAAAC  
 901 GTCGAGGCCG CGATTAAAAAT CCAACATGGA TGCTGATTG TATGGGTATA AATGGGCTCG  
 961 CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG TATGGGAAGC CCGATGCGCC  
 1021 AGAGTTGTTT CTGAAACATG GCAAAAGGTAG CGTTGCCAAT GATGTTACAG ATGAGATGGT  
 1081 CAGACTAAAC TGGCTGACGG AATTATGCC TCTTCCGACC ATCAAGCATT TTATCCGTAC  
 1141 TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGAA AAAACAGCAT TCCAGGTATT  
 1201 AGAAGAATAT CCTGATTCAAG GTGAAAATAT TGTTGATGCG CTGGCAGTGT TCCCTGCGCCG  
 1261 GTTGCATTGCG ATTCCCTGTT GTAAATTGTC TTTAACAGC GATCGCGTAT TTCTGCTCGC  
 1321 TCAGGCGCAA TCACGAATGA ATAACGGTTT GGTTGATGCG AGTGTATTTG ATGACGAGCG  
 1381 TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT AAACTTTGC CATTCTCAC  
 1441 GGATTCACTGC GTCACACTCATG GTGATTCTC ACTTGATAAC CTTATTTTG ACGAGGGGAA  
 1501 ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA GACCGATAACC AGGATCTTGC  
 1561 CATCCTATGG AACTGCCCTCG GTGAGTTTTC TCCTTCATTA CAGAAACGGC TTTTCAAAA  
 1621 ATATGGTATT GATAATCCCG ATATGAATAA ATTGAGTTT CATTGATGC TCGATGAGTT  
 1681 TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA GATTGGGCC CGTCCCACTG  
 1741 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTT TTCTGCGCGT  
 1801 AATCTGCTGC TTGCAAACAA AAAAACACC GCTACCAGCG GTGGTTGTT TGCGGGATCA  
 1861 AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC  
 1921 TGTTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC  
 1981 ATACCTCGCT CTGCTAATCC TGTTACAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT  
 2041 TACCGGGTTG GACTCAAGAC GATAAGTACCC GGATAAGGCG CAGCGGTCGG GCTGAACGGG  
 2101 GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACTACA  
 2161 GCGTGAGCTA TGAGAAAAGCG CCACCGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT  
 2221 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCTGGTA  
 2281 TCTTTATAGT CCTGTCGGGT TTGCGCACCT CTGACTTGAG CGTCGATTT TGTGATGCTC  
 2341 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACCGCG GCCTTTTAC GGTTCTGGC  
 2401 CTTTTGCTGG CCTTTTGTCTC ACATGTTCTT TCCTGCGTTA TCCCCGTATT CTGTTGATAA  
 2461 CCGTATTACG GCTAGCATGG ATCTCGGGGA CGTCTAACTA CTAAGCGAGA GTAGGAAACT  
 2521 GCCAGGCATC AAAATAAAACG AAAGGCTCAG TCGGAAGACT GGGCCTTCG TTTTATCTGT  
 2581 TGTTTGTGG TGAAACGCTCT CCTGAGTAGG ACAAAATCCGC CGGGAGCGGA TTTGAACGTT  
 2641 GTGAAGCAAC GGCCCCGGAGG GTGGCGGGCA GGACGCCCGC CATAAACTGC CAGGCATCAA  
 2701 ACTAAGCAGA AGGCCATC

FIGURE 1(B)

**Figure 7A: Cloning Sites of the Entry Vector pENTR3C (reading frame C)**

Int	attL1	DraI	XmnI	SalI	BamHI
TTG TAC AAA AAA GCA GGC TCT TTA AAG GAA CCA ATT CAG TCG ACT GGA TCC GGT					
AAC ATG TTT TTT CGT CCG AGA AAT TTC CTT GGT TAA GTC AGC TGA CCT AGG CCA					
Leu Tyr Lys Lys Ala Gly Ser Leu Lys Glu Pro Ile Gln Ser Thr Gly Ser Gly					

KpnI	EcoRI	PvuI	EcoRI	NotI	XbaI	EcoRV	XbaI
ACC GAA TTC GAT CGC-- ccdB	--G	AAT TCG CGG CCG CAC TCG AGA TAT CTA					
TGG CTT AAG CTA GCG	C	TTA AGC GCC GGC GTG AGC TCT ATA GAT					
Thr Glu Phe		Asn Ser Arg Pro His Ser Arg Tyr Leu					

attL2	Int
GAC CCA GCT TTC TTG TAC AAA G	
CTG GGT CGA AAG AAC ATG TTT C	
Asp Pro Ala Phe Leu Tyr Lys	

15/240

## pENTR3C 2723 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
327..632	ccdB
661..760	attL2
883..1692	KmR
1797..2370	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTC ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAACATTG TACAAAAAAAG CAGGCTCTT AAAGGAACCA  
 181 ATTCAAGTCGA CTGGATCCCG TACCGAATTG GATCGCTTAC TAAAAGCCAG ATAACAGTAT  
 241 GCGTATTTCG GCGCTGATTG TTGCGGTATA AGAATATATA CTGATATGTA TACCCGAAGT  
 301 ATGTCAAAAAA GAGGTGTGCT TCTAGAATGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA  
 361 GCCGTTATCG TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCC GGGCGACGGA  
 421 TGGTGATCCC CCTGGCCAGT GCACGCTCTGC TGTAGATCAA AGTCTCCGT GAACTTTACC  
 481 CGGTGGTGC TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC  
 541 CGGTCTCCGT TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA  
 601 ACGCCATTAA CCTGATGTTG TGGGAATAT AGAATTGCG GCGCACTCG AGATATCTAG  
 661 ACCCAGCTT CTTGTACAAA GTTGGCATTA TAAGAAAGCA TTGTTATCA ATTTGTTGCA  
 721 ACGAACAGGT CACTATCACT CAAATAAAA TCATTATTTG CCATCCAGCT GCAGCTCTGG  
 781 CCCGTGTCTC AAAATCTCTG ATGTTACATT GCACAAGATA AAAATATATC ATCATGAACA  
 841 ATAAAACGTG CTGCTTACAT AAACAGTAAT ACAAGGGGTG TTATGAGCCA TATTCAACGG  
 901 GAAACGTGCA GGCGCGATT AAATCCAAC ATGGATGCTG ATTTATATGG GTATAAATGG  
 961 GCTCGCGATA ATGTCGGCA ATCAGGTGCG ACAATCTATC GCTTGTATGG GAAGCCCGAT  
 1021 GCGCCAGAGT TGTTTCTGAA ACATGGAAA GTAGCGTTG CCAATGATGT TACAGATGAG  
 1081 ATGGTCAGAC TAAACTGGCT GACGGAATTG ATGCTCTTC CGACCATCAA GCATTTATC  
 1141 CGTACTCCTG ATGATGCATG GTTACTCACC ACTGCGATCC CCGGAAAAAC AGCATTCCAG  
 1201 GTATTAGAAG AATATCCTGA TTCAGGTGAA AATATTGTT ATGCGCTGGC AGTGTTCCTG  
 1261 CGCCGGTTGC ATTGATTCC TGTTTGTAAAT TGTCCTTTA ACAGCGATCG CGTATTTCGT  
 1321 CTCGCTCAGG CGCAATCAAG AATGAATAAC GGTTGGTTG ATGCGAGTGA TTTTGATGAC  
 1381 GAGCGTAATG GCTGGCTGT TGAAACAAGTC TGGAAAGAAA TGATAAAACT TTTGCCATT  
 1441 TCACCGGATT CAGTCGTAC TCATGGTGAT TTCTCACTTG ATAACCTTAT TTTGACGAG  
 1501 GGGAAATTAA TAGGTTGTAT TGATGTTGGA CGAGTCGGAA TCGCAGACCG ATACCAAGGAT  
 1561 CTTGCCATCC TATGGAACCTG CCTCGGTGAG TTTTCTCCTT CATTACAGAA ACGGCTTTT  
 1621 CAAAAATATG GTATTGATAA TCCTGATATG AATAAAATTG AGTTTCATTT GATGCTCGAT  
 1681 GAGTTTTCT AATCAGAATT GGTTAATTGG TTGTAACATT ATTCAAGATTG GGCCCCGTTC  
 1741 CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTGAGATCC TTTTTTCTG  
 1801 CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGGTGT TTGTTTGC  
 1861 GATCAAGAGC TACCAACTCT TTTCCGAAG GTAACTGGCT TCAGCAGAGC GCAGATAACCA  
 1921 AATACTGTTTCTTCTAGTGTGAA GCCGTAGTTA GCCCACCACT TCAAGAACCTC TGAGCACCG  
 1981 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG  
 2041 TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA  
 2101 ACGGGGGGTT CGTGACACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC  
 2161 CTACAGCGTC AGCTATGAGA AAGGCCACAG CTTCCCGAAG GGAGAAAGC GGACAGGTAT  
 2221 CCGGTAAGCG GCAGGGTGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC  
 2281 TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA  
 2341 TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGCCCTT TTTACGGTTC  
 2401 CTGGCCCTTTT GCTGGCCCTT TGCTCACATG TTCTTCTGCT CGTTATCCCC TGATTCTGTG  
 2461 GATAACCGTA TTACCGCTAG CATGGATCTC GGGGACGTCT AACTACTAAG CGAGAGTAGG  
 2521 GAACTGCCAG GCATCAAATA AAACGAAAGG CTCAGTCGGA AGACTGGGCC TTTCGTTTAA  
 2581 TCTGTTGTTT GTCGGTGAAAC GCTCTCCTGA GTAGGACAAA TCCGCCGGGA GCGGATTGAA  
 2641 ACGTGTGAA GCAACGGCCC GGAGGGTGGC GGGCAGGACG CCCGCCATAA ACTGCCAGGC  
 2701 ATCAAACATAA CGAGAAGGCC ATC

FIGURE 12B

16/240

**Figure 13A: Cloning Sites of the Entry Vector pENTR4**

KpnI EcoRI	EcoRI	NotI	XhoI	EcoRV	XbaI
<u>TAC</u> CGA ATT C-- ccdB	--G <u>AAT</u> TCG <u>CTG</u> CCG CAC <u>TCG</u> AGA TAT <u>CTA</u> GAC CCA GCT				
ATG GCT TAA G	C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT CGA				
Tyr Arg Ile	Asn Ser Arg Pro His Ser Arg Tyr Leu Asp Pro Ala				

Int attL2  
 TTC TTG TAC AAA G  
 AAG AAC ATG TTT C  
 Phe Leu Tyr Lys

17/240

## pENTR4 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

1 CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCGT TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTCCAC CATGGAAACC  
 181 AATTCACTCG ACTGGATCCG GTACCGAATT CGCTTACTAA AAGCCAGATA ACAGTATGCG  
 241 TATTTGCGCG CTGATTTTG CGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG  
 301 TCAAAAAGAG GTGTGCTTCT AGAATGCAGT TTAAGGTTA CACCTATAAA AGAGAGAGCC  
 361 GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCGGG CGACGGATGG  
 421 TGATCCCCCT GGCCAGTGC CGTCTGCTG CAGATAAAAGT CTCCCGTGA CTTTACCCGG  
 481 TGGTGCATAT CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGCC AGTGTGCCGG  
 541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAACG  
 601 CCATTAACCT GATGTTCTGG GGAATATAGA ATTGCGGGCC GCACTCGAGA TATCTAGACC  
 661 CAGCTTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTGCAACG  
 721 AACAGGTACAC TATCAGTCAA AATAAAATCA TTATTTGCCA TCCAGCTGCA GCTCTGGCCC  
 781 GTGTCTCAAAT ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA  
 841 AAACGTGCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA  
 901 ACGTCGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGTA TAAATGGGCT  
 961 CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGCT TGTATGGAA GCCGATGCG  
 1021 CCAGAGTTGT TTCTGAAACA TGGCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG  
 1081 GTCAGACTAA ACTGGCTGAC GGAATTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGT  
 1141 ACTCCTGGTG ATGCATGGTT ACTCACCACT GCGATCCCCG GAAAACAGC ATTCCAGGTA  
 1201 TTAGAAGAAT ATCCGTATTC AGGTGAAAAT ATTGTTGATG CGCTGGCAGT GTTCTGCGC  
 1261 CGGTTGCATT CGATTCCTGT TTGTAATTGT CCTTTTAACA GCGATCGCGT ATTTCGTCTC  
 1321 GCTCAGGGCGC AATCACGAAT GAATAACCGT TTGGTTGATC CGAGTGAATT TGATGACGAG  
 1381 CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAAGAAATGC ATAAACTTTT GCCATTCTCA  
 1441 CCGGATTCAAG TCGTCACTCA TGGTGAATTCA TCACTTGATA ACCTTATTT TGACGAGGGG  
 1501 AAATTAATAG GTTGTATTGA TGTTGGACGA GTCGGAATCG CAGACCGATA CCAGGATCTT  
 1561 GCCATCCTAT GGAACTGCC CGGTGAGTTT TCTCCTTCTAT TACAGAAACG GCTTTTCAA  
 1621 AAATATGGTA TTGATAATCC TGATATGAAT AAATTGCACT TTCATTGAT GCTCGATGAG  
 1681 TTTTCTAAT CAGAATTGGT TAATTGTTG TAAACATTATT CAGATTGGC CCCGTTCCAC  
 1741 TGAGCGTCAG ACCCCGTAGA AAAGATCAA GGATCTTCTT GAGATCTTTT TTTCTGCGC  
 1801 GTAATCTGCT GTTGCAAAC AAAAAGACCA CCGCTACCAAG CGGTGGTTG TTTGCCGGAT  
 1861 CAAGAGCTAC CAACTTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGA GATACCAAT  
 1921 ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTC AAGAACTCTGT AGCACCGCCT  
 1981 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT  
 2041 CTTACCGGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG  
 2101 GGGGGTTCTGT GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA  
 2161 CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG  
 2221 GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCCTGG  
 2281 TATCTTTATA GTCCCTGCTGG GTTTCGCCAC CTCTGACTTG AGCGTGGATT TTTGTGATGC  
 2341 TCGTCAGGGGG GGGGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTT ACGGTTCCGT  
 2401 GCCTTTTGTGCT GGCCTTTGTC TCACATGTT TCACATGTT TATCCCTGTA TTCTGTGGAT  
 2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAACGCA GAGTAGGGAA  
 2521 CTGCCAGGCA TCAAATAAAA CGAAAGGCTC AGTCGGAAGA CTGGGCCCTT CGTTTTATCT  
 2581 GTTGTGGTCT GGTGAACGCT CTCCTGAGTA GGACAAATCC GCCGGGAGCG GATTGAAACG  
 2641 TTGTGAAGCA ACGGCCCGGA GGGTGGCGGG CAGGACGCC GGCATAAACT GCCAGGCATC  
 2701 AACTAAGCA GAAGGCCATC

FIGURE 13B

18/240

Figure 14A: Cloning sites of the Entry Vector pENTR5

Int att L1 Nde I Xba I Sst I  
 Tgg tac aaa aaa gca ggc tt cat atg gaa atc aat tca gtc  
 Atc atg ttt ttt cgt ccg aha gta ttc cct tgg tta agt cag  
 Leu Tyr Lys Lys Ala Gly Phe His Met Gly Thr Asn Ser Val

Bam HI Kpn I Eco RI Eco RI  
 gac tgg atc cgg tac cga att cgc --- Death --- agt att cgc  
 ctg acc tag ggc atg gct taa gcg --- (ccdB) --- tct taa gcg.  
 Asp Trp Ile Arg Tyr Arg Ile

Nsi I Xba I Eco RI Xba I Int att L2  
 bge cgc act cga gat atc tag acc cag ctt tcc xgg aga aag  
 ccg gcg tga gct cta tag atc tgg gtc gaa aga aca cgt tcc

19/240

## pENTR5 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTC ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTTCA TATGGGAACC  
 181 AATTCACTCG ACTGGATCCG GTACCGAATT CGCTTACTAA AAGCCAGATA ACAGTATGCC  
 241 TATTTGCGCG CTGATTTTG CGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG  
 301 TCACAAAGAG GTGTGCTTCT AGAATGCACT TTAAGGTTA CACCTATAAA AGAGAGAGCC  
 361 GTTATCGTCT GTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCGGG CGACGGATGG  
 421 TGATCCCCCT GGCCAGTGCAG CGTCTGCTGT CAGATAAAAGT CTCCCGTCAA CTTTACCCGG  
 481 TGGTGCATAT CGGGGATGAA AGCTGGCGCA TGATGACCA CGATATGCC AGTGTGCCGG  
 541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAACG  
 601 CCATTAACCT GATGTTCTGG GGAATATAGA ATTGCGGGCC GCACTCGAGA TATCTAGACC  
 661 CAGCTTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTGCAACG  
 721 AACAGGTCAAC TATCAGTCAA AATAAAATCA TTATTTGCCA TCCAGCTGCA GCTCTGGCCC  
 781 GTGTCTAAA ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAAACAATA  
 841 AAACGTCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA  
 901 ACGTCGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGTA TAAATGGCT  
 961 CGCGATAATG TCGGGCAATC AGGTGGGACA ATCTATCGCT TGTATGGAA GCCCGATGCC  
 1021 CCAGAGTTGT TTCTGAAACA TGGCAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG  
 1081 GTCAGACTAA ACTGGCTGAC GGAATTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGT  
 1141 ACTCTGATG ATGCATGGTT ACTCACCAC GCGATCCCCG GAAAACAGC ATTCCAGGTA  
 1201 TTAGAAGAAT ATCCTGATTC AGGTGAAAT ATTGTTGATG CGCTGGCAGT GTTCTGCCG  
 1261 CGGTTGCATT CGATTCCTGT TTGTAATTGT CCTTTAACCA GCGATCGCGT ATTTCTGCTC  
 1321 GCTCAGGCGC AATCACGAAT GAATAACGGT TTGGTTGATG CGAGTGAATT TGATGACGAG  
 1381 CGTAATGGCT GGCGCTGTTGA ACAAGTCTGG AAAGAAATGC ATAAACTTTT GCCATTCTCA  
 1441 CGGGATTCACTCG TCGTCACTCA TGGTGATTTC TCACCTGATA ACCTTATTT TGACGAGGGG  
 1501 AAATTAATAG GTTGTATTGA TGTTGGACGA GTGCGGAATCG CAGACCGATA CCAGGATCTT  
 1561 GCCATCCTAT GGAACCTGCT CGGTGAGTTT TCTCCTTCAT TACAGAAACG GCTTTTTCAA  
 1621 AAATATGGTA TTGATAATCC TGATATGAAT AAATTGCACT TTCATTGAT GCTCGATGAG  
 1681 TTTTCTAAT CAGAATTGTT TAATTGGTT TAACATTATT CAGATTGGC CCCGTTCCAC  
 1741 TGAGCGTCAG ACCCGTAGA AAAGATCAAA GGATCTTCIT GAGATCCTTT TTTTCTGCC  
 1801 GTAATCTGCT GCTTGCAAAAC AAAAAGACCA CCGCTTACAG CGGTGGTTTG TTTGCCGGAT  
 1861 CAAGAGCTAC CAACTCTTT TCCGAAGGTA ACTGGCTTCA GCAGAGGCCA GATACCAAAT  
 1921 ACTGTTCTTC TAGTGTAGGC GTAGTGTAGGC CACCACTTCA AGAACCTCTGT AGCACCGCCT  
 1981 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTTG  
 2041 CTTACGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGTC GGGCTGAACG  
 2101 GGGGGTTCGT GCACACAGGC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA  
 2161 CAGCGTGAAG TATGAGAAAG CGCCACGCTT CCCGAAGGGGA GAAAGGCCA CAGGTATCCG  
 2221 GAAAGCGGCA GGGTGGAAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCCTGG  
 2281 TATCTTTATA GTCTGTGCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC  
 2341 TCGTCAGGGG GGCAGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTT ACGGTTCTG  
 2401 GCCCTTTGCT GGCCTTTTGC TCACATGTTT TTTCCTGCGT TATCCCCCTGA TTCTGTGGAT  
 2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAACGGA GAGTAGGGAA  
 2521 CTGCCAGGCA TCGAATAAA CGAAAGGCTC AGTCGGAAGA CTGGGCCTTT CGTTTTATCT  
 2581 GTTGTGTC GGTGAACGCT CTCCGTAGTA GGACAAATCC GCGGGGAGCG GATTTGAACG  
 2641 TTGTGAAGCA ACGGCCCCGA GGGTGGCGGG CAGGACGCC GCGATAAACT GCCAGGCATC  
 2701 AACTAAAGCA GAAGGCCATC

Figure 14B

20/240

Figure 15A: Cloning sites of the Entry Vector pENTR6

Int att L1 Sph I Kpn I Xba I Sal I  
 --- ttp tac aaa aaa gca ggc tgc atg cga acc aat tca gtc  
acc aac cgt ttt cgt cgg agg tac gct tgg tta agt cag  
 Leu Tyr Lys Lys Ala Gly Cys Met Arg Thr Asn Ser Val

BamH I Kpn I EcoRI EcoRI  
 gac tgg atc cgg tac cga att cgc --- Death --- agt att cgc  
tgg acc tag gct atg gct taa gct --- (codB) --- tct taa gct  
 Asp Trp Ile Arg Tyr Arg Ile

Not Xba I EcoR I Xba I Int att L2  
 bgc cgc act cga gat atc tag acc cag ctt tgc aca gag ---  
ccg ccg tga gct cta tag atc tgg gtc gaa aga aca tgt tcc ---

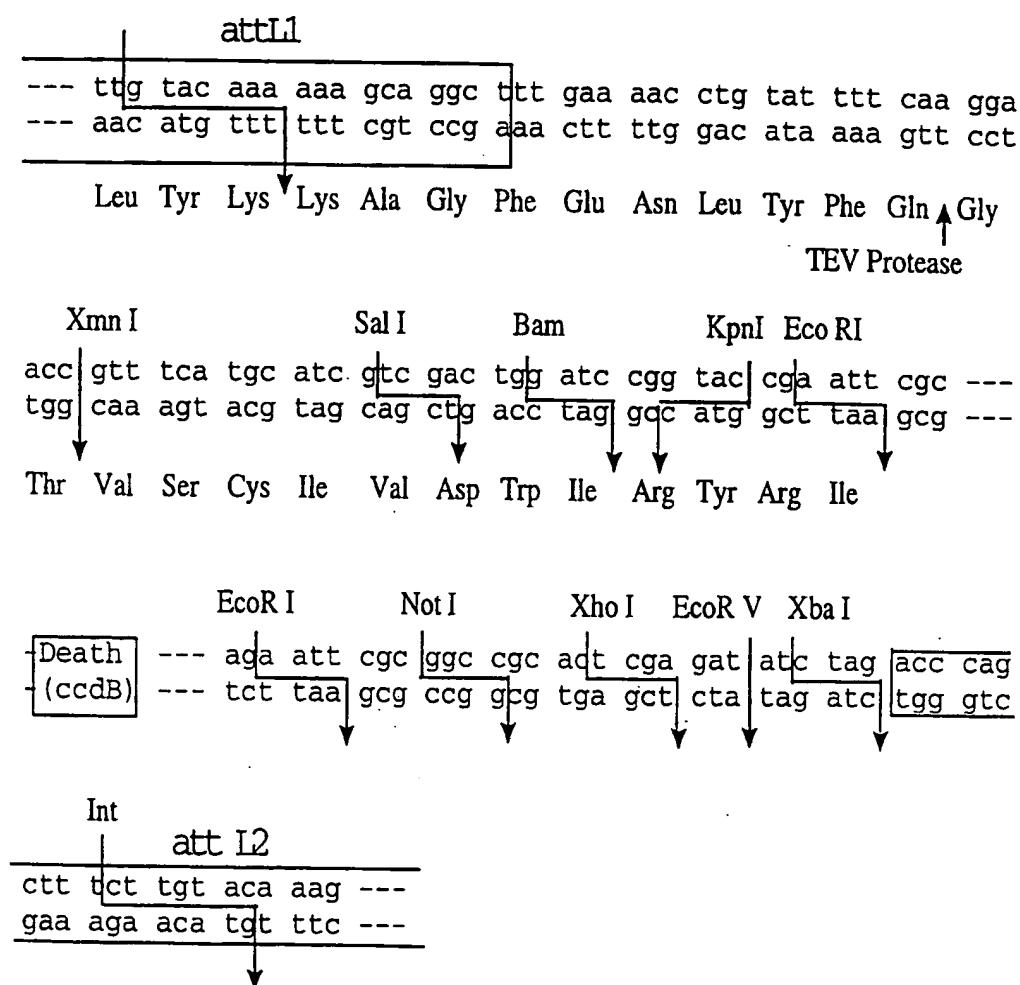
21/240

## pENTR6 2717 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

1 CTGACGGATG GCCTTTTTCG GTTTCTACAA ACTCTTCCGT TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAATTGTTG TACAAAAAAG CAGGCTGCAT GCGAACCAAT  
 181 TCAGTCGACT GGATCCGGTA CGGAATTCCGC TTACTAAAAG CCAGATAACA GTATGCGTAT  
 241 TTGCGCGCTG ATTTTGCGG TATAAGAATA TATACTGATA TGATACCCG AAGTATGTCA  
 301 AAAAGAGGTG TGCTTCTAGA ATGCGAGTTA AGGTTTACAC CTATAAAAGA GAGAGCCGTT  
 361 ATCGTCTGTT TGTGGATGTA CAGAGTGTATA TTATTGACAC GCCCCGGCGA CGGATGGTGA  
 421 TCCCCCTGGC CAGTGCACGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG  
 481 TGCATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA TATGCCAGT GTGCCGGTCT  
 541 CCGTTATCGG GGAAGAAGTG GCTGATCTA GCCACCGCGA AAATGACATC AAAAACGCCA  
 601 TTAACCTGAT GTTCTGGGA ATATAGAATT CGCGGCGCGA CTCGAGATAT CTAGACCCAG  
 661 CTTTCTTGTA CAAAGTTGGC ATTATAAGAA AGCATTGCTT ATCAATTGTT TGCAACGAAC  
 721 AGGTCACTAT CAGTCAAAAT AAAATCATTA TTGCCCATTG AGCTGCAGCT CTGGCCCGTG  
 781 TCTCAAATC TCTGATGTTA CATTGACCAA GATAAAAATA TATCATCATG AACATAAAA  
 841 CTGTCGCTT ACATAAACAG TAATACAAGG GGTGTTATGA GCCATATTCA ACGGGAAACG  
 901 TCGAGGCCGC GATTAAATTG CAACATGGAT GCTGATTAT ATGGGTATAA ATGGGCTCGC  
 961 GATAATGTCG GGCAATCAGG TGCACAAATC TATCGCTTGT ATGGGAAGCC CGATGCGCCA  
 1021 GAGTTGTTTC TGAAAACATGG CAAAGGTAGC GTTGCAATG ATGTTACAGA TGAGATGGTC  
 1081 AGACTAAACT GGCTGACCGA ATTTATGCTT CTTCGACCA TCAAGCATTT TATCCGTACT  
 1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCCGGAA AAACAGCATT CCAGGTATTAA  
 1201 GAAGAAATATC CTGATTCAAGG TGAAAATATT GTTGATGCGC TGGCAGTGT CCTGGCCCGG  
 1261 TTGCAATTGCA TTCTGTTTG TAATTGCTT TTTAACAGCG ATCGCGTATT TCGTCTCGCT  
 1321 CAGGCGCAAT CACGAATGAA TAACGGTTTG GTTGATGCGA GTGATTTGA TGACGGAGCGT  
 1381 AATGGCTGGC CTGTTGAACA AGTCTGGAAA GAAATGCATA AACTTTGCC ATTCTCACCG  
 1441 GATTCACTCG TCACTCATGG TGATTCTCA CTTGATAACC TTATTTTGAA CGAGGGAAA  
 1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATACCA GGATCTTGCC  
 1561 ATCCTATGGA ACTGCCTCGG TGAGTTTCT CCTTCATTAC AGAAACGGCT TTTTCAAAA  
 1621 TATGGTATTG ATAATCTGA TATGAATAA TTGAGTTTC ATTTGATGCT CGATGAGTTT  
 1681 TTCTAATCAG AATTGGTAA TTGGTTGAA CATTATTCAAGG ATTGGGCCCG GTTCACTGAA  
 1741 GCGTCAGACC CCGTAGAAAA GATCAAAAGGA TCTTCTTGAG ATCCCTTTT TCTGCGCGTA  
 1801 ATCTGCTGCT TGCAAAACAAA AAAACACCG CTACCAAGCGG TGGTTGTTT GCCGATCAA  
 1861 GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA GAGCGCAGAT ACCAAATACT  
 1921 GTTCTCTAG GTAGCCGTAA GTTACGCCAC CACTTCAGA ACTCTGTAAC ACCGCCTACA  
 1981 ACCCTCGCTC TGCTAATCTT GTTACCAAGT GCTGCTGCCA GTGGCGATAA GTCGTGTCTT  
 2041 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTGGG CTGAACGGGG  
 2101 GTTCTGCA CACAGCCCG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG  
 2161 CGTGAGCTAT GAGAAAGCCG CACGGCTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA  
 2221 AGCGGCAGGG TCGGAACAGG AGAGCGCAG AGGGAGCTC CAGGGGGAAA CGCCTGGTAT  
 2281 CTTTATAGTC CTGTCGGTTT TCGCCACCTC TGACTTGAGC GTGATTGTTT GTGATGCTCG  
 2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCC CTTTTTACG GTTCCCTGGCC  
 2401 TTTTGCTGGC CTTTGTCTA CATGTTCTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC  
 2461 CGTATTACCG CTAGCATGGA TCTCGGGAC GTCTAACTAC TAAGCGAGAG TAGGAAACTG  
 2521 CCAGGCATCA AATAAAACGA AAGGCTAGT CGGAAGACTG GGCCTTCGT TTTATCTGTT  
 2581 GTTTGTCGGT GAACGCTCTC CTGAGTAGGA CAATCCGCC GGGAGCGGAT TTGAACGTTG  
 2641 TGAAGCAACG GCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGCATCAAA  
 2701 CTAAGCAGAA GGCCATC

FIGURE 15B

**Figure 16A: Cloning sites of the Entry Vector pENTR7**

23/240

## pENTR7 2738 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTT ATTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAAG CAGGCTTGA AAACCTGTAT  
 181 TTCAAGGAA CCGTTCATG CATCGTCGAC TGGATCCGGT ACCGAATTG CTTACTAAA  
 241 GCCAGATAAAC AGTATGCGTA TTTGCGCGCT GATTTTGC GTATAAGAAT ATATACTGAT  
 301 ATGTATACCC GAAGTATGTC AAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA  
 361 CCTATAAAAAG AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA  
 421 CGCCCGGGCG ACGGATAGTG ATCCCCCTGG CCAGTGCACG TCTGCTGCA GATAAAGTCT  
 481 CCCGTGAAC TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG  
 541 ATATGGCCAG TGTGCCGGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG  
 601 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTTCTGGGG AATATAGAAT TCGCGGCGC  
 661 ACTCGAGATA TCTAGACCA GCTTCTTGT ACAAAAGTTGG CATTATAAGA AAGCATTGCT  
 721 TATCAATTTG TTGCAACGAA CAGGTCACTA TCAGTCAAA TAAAATCATT ATTTGCCATC  
 781 CAGCTGCAGC TCTGGCCCGT GTCTAAAAT CTCTGATGTT ACATTGCACA AGATAAAAAT  
 841 ATATCATCAT GAACAATAAA ACTGCTGCT TACATAAAACA GTAATACAAG GGGTGTATG  
 901 AGCCATATTC AACGGGAAAC GTCGAGGCCG CGATTAATT CCACATGGA TGCTGATTTA  
 961 TATGGGTATA AATGGGCTCG CGATAATGTC GGGCAATCG GTGCGACAA CTATCGCTTG  
 1021 TATGGGAAGC CCGATGCGCC AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT  
 1081 GATGTTACAG ATGAGATGGT CAGACTAAAC TGGCTGACGG AATTTATGCC TCTTCCGACC  
 1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGA  
 1201 AAAACAGCAT TCCAGGTATT AGAAGAATAT CCTGAATTGAC GTGAAAATAT TGTTGATGCG  
 1261 CTGGCAGTGT TCCGTGCCCG GTTGCAATTG ATTCCCTGTTT GTAATTGTC TTTAACAGC  
 1321 GATCGCGTAT TTCGTCTCGC TCAGCGCAA TCACGAATGA ATAACGGTTT GGTTGATGCG  
 1381 AGTGTATTTG ATGACGAGCG TAATGGCTGG CCTGTTGAAAC AAGTCTGGAA AGAAATGCAT  
 1441 AAACCTTTGC CATTCTCACCG GGATTCACTG GTCACTCATG GTGATTTCTC ACTTGATAAC  
 1501 CTTATTTTG ACGAGGGGAA ATTAATAGGT TGTATTGATG TTGGACCGAGT CGGAATCGCA  
 1561 GACCGATACC AGGATCTTGC CATCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA  
 1621 CAGAAACGGC TTTTCAAAA ATATGGTATT GATAATCTG ATATGAATAA ATTGCAGTTT  
 1681 CATTGATGC TCGATGAGTT TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA  
 1741 GATTGGGCC CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTCTTGA  
 1801 GATCCTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCAAC GCTACCAGCG  
 1861 GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC  
 1921 AGAGCGCAGA TACCAAATAC TGTTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG  
 1981 AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC  
 2041 AGTGGCGATA AGTCGTGCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG  
 2101 CAGCGGTGCG GCTGAACGGG GGGTCTGTC ACACAGCCCA GCTTGGAGCG AACGACCTAC  
 2161 ACCGAACCTGA GATACTACA CGCTGAGCTA TGAGAAAGCG CCACCGCTCC CGAACGGAGA  
 2221 AAGGGCGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCCAC GAGGGAGCTT  
 2281 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTGCCCACCT CTGACTTGAG  
 2341 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAAACGC CAGCAACCGC  
 2401 GCCTTTTAC GGTTCTGGC CTTTGTCTGG CCTTTTGTCTC ACATGTTCTT TCCTGCGTTA  
 2461 TCCCCTGATT CTGTGGATAA CGTATTACG GCTAGCATGG ATCTCGGGGA CGTCTAACTA  
 2521 CTAAGCGAGA GTAGGGAACT GCCAGGCATC AAATAAAAG AAAGGCTCAG TCGGAAGACT  
 2581 GGGCCTTCG TTTATCTGT TGTTGTCTGG TGAACGCTCT CCTGAGTAGG ACAAAATCCGC  
 2641 CGGGAGCGGA TTTGAACGTT GTGAAGCAAC GGCCCGGAGG GTGGCGGGCA GGACGCCCCG  
 2701 CATAAAACTGC CAGGCATCAA ACTAAGCAGA AGGCCATC

FIGURE 16B

24/240

Figure 17A: Cloning Sites of the Entry Vector pETURB

Int attL

tac aaa aaa gca ggc ttt gaa aac ctg tat ttt caa gya  
ttt cgt ccg aaa ctt ttg gac ata aaa gtt cct

Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln, Gly

TEV Protease

NcoI AA II S<sub>a</sub>I BamH I KpnI EcoRI

atc atg gac cta gtc gac tgt atc cgg tac cda att cgc ---  
tgg tac ctg gat cag ctg acc tag gtc atg gct taa gcg ---

Thr Met Asp Leu Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI Not I XbaI EcoRI XbaI attL

Death --- aga att cgc ggc cgc act cga gat atc tag acc cag  
--- tct taa gcg ccg ccg tga gct cta tag atc tgg gtc

Int

cet tct xgt aca aaq ---  
gaa aga aca tgt ttc ---

25/240

## pENTR8 2735 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
339..644	ccdB
673..772	attL2
895..1704	KmR
1809..2382	ori

1 CTGACGGATG GCCTTTTGTG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTC ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTATAAT GCCAACATTG TACAAAAAAAG CAGGCTTGA AAACCTGTAT  
 181 TTTCAGGAA CCATGGACCT AGTCGACTGG ATCCGGTACCGA ATTTCGCTT ACTAAAAGCC  
 241 AGATAAACAGT ATGCGTATTT GCGCGCTGAT TTTTGCCTTA TAAGAATATA TACTGATATG  
 301 TATACCCGAA GTATGTCAA AAGAGGTGTG CTCTAGAAT GCAGTTAACCT GTTTACACCT  
 361 ATAAAAGAGA GAGCGTTAT CGTCTGTTG TGATGTACA GAGTGTATT ATTGACACGC  
 421 CCGGGCGACG GATAGTGATC CCCCTGGCCA GTGCACGCTC GCTGTCAGAT AAAGTCTCCC  
 481 GTGAACCTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA  
 541 TGGCCAGTGT GCGGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA  
 601 ATGACATCAA AAACGCCATT AACCTGTGTT TCTGGGAAT ATAGAATTG CGGCCGACT  
 661 CGAGATATCT AGACCCAGCT TTCTTGTACA AGTGGCAT TATAAGAAAG CATTGCTTAT  
 721 CAATTGTTG CAACGAACAG GTCACTATCA GTCAAAATAA AATCATTATT TGCCATCCAG  
 781 CTGCAGCTCT GGCCCGTGTG TCAAAATCTC TGATGTACA TTGCACAAAGA TAAAATATA  
 841 TCATCATGAA CAATAAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC  
 901 CATATTCAAC GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT  
 961 GGGTATAAAAT GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGTAT  
 1021 GGGAAAGCCCG ATGCGCCAGA GTTGTTCCTG AAACATGGCA AAGGTAGCGT TGCCAAATGAT  
 1081 GTTACAGATG AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC  
 1141 AAGCATTTC TCCGTACTCC TGATGTGCA TGGTTACTCA CCACTGCGAT CCCCCGAAAA  
 1201 ACAGCATTCC AGGTATTAGA AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG  
 1261 GCACTGTCCTC TGCGCCGGTT GCATTGATT CCTGTTGTA ATTGCTCTTT TAACAGCGAT  
 1321 CGCGTATTTG GTCTCGCTCA GGCGCAATCA CGAATGAATA ACGGTTGGT TGATGCGAGT  
 1381 GATTTTGATG ACGAGCGTAA TGGCTGGCCT GTGAACAAG TCTGGAAAGA AATGCATAAAA  
 1441 CTTTGCCAT TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT  
 1501 ATTTTGACG AGGGGAAATT AATAGGTGTG ATTGATGTG GACGAGTCGG AATCGCAGAC  
 1561 CGATACCAGG ATCTTGCCAT CCTATGGAAC TGCTCGGGTG AGTTTCTCC TTCATTACAG  
 1621 AAACGGCTTT TTCAAAATAA TGGTATTGAT AATCCTGATA TGAATAAAATT GCAGTTTCAT  
 1681 TTGATGCTCG ATGAGTTTT CTAATCAGAA TTGGTTAATT GTGTGAACA TTATTCAAGAT  
 1741 TGGGCCCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAAGGATC TTCTTGAGAT  
 1801 CCTTTTTTTC TGCGCGTAAT CTGCTGTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG  
 1861 GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTGG CTTCAGCAGA  
 1921 GCGCAGATAC CAAATACTGT TCTCTAGTG TAGCGCTAGT TAGGCCACCA CTTCAAGAAC  
 1981 TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCACTGGC TGCTGCCAGT  
 2041 GCGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG  
 2101 CGCTGGGGCT GAACGGGGGG TTGCTGCACA CAGCCCGACT TGGAGCGAAC GACCTACACC  
 2161 GAACGTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG  
 2221 GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA  
 2281 GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT  
 2341 CGATTTTGTG GATGCTCGTC AGGGGGGGGG AGCCTATGGA AAAACGCCAG CAACGCGGCC  
 2401 TTTCGACGGT TCCTGGCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC TGCGTTATCC  
 2461 CCTGATTCTG TGGATAACCG TATTACCGCT AGCATGGATC TCAGGGACGT CTAACTACTA  
 2521 AGCGAGAGTA GGGAACTGCC AGGCATCAA TAAAACGAAA GGCTCAGTCG GAAGACTGGG  
 2581 CCTTTGCTTT TATCTGTGTT TTGTCGGTGA ACGCTCTCT GAGTAGGACA AATCCGCCGG  
 2641 GAGCGGATTT GAACGTTGTG AAGCAACGGC CGGGAGGGTG GCAGGGCAGGA CGCCCGCCAT  
 2701 AAACTGCCAG GCATCAAAC AAGCAGAAGG CCATC

FIGURE 17B

76/240

Figure 18A: Cloning sites of the *EntY* Vector pENTR29

Int attL1

ttg tac aaa aaa gca ggc ttt gaa aac ctg tat ttt caa gga  
 --- aca atg tkt ttt cgt ccg aca ctt ttg gac ata aaa gtt cct  
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln Gly  
 TEV protease

NdeI BglII SalI BamHI KpnI EcoRI  
 cat atg aca tct gtc gac tgg atc cgg tac cga att cgc ---  
 gta tac tct aca cag cgg acc tag gtc atg gct taa gcg ---  
 His Met Arg Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NheI XbaI EcoRI NheI attL2  
 Deletion --- aga att cgc ggc egc act cga gat att tag acc cag  
 --- tct taa gcg eeg gcg tga gct cta tag atc tgg gtc

Int

ctc tct tgg aca aca ---  
 gaa aga aca tgt tcc ---

27/240

## pENTR9 2735 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
339..644	ccdB
673..772	attL2
895..1704	KmR
1809..2382	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTGAA AACCTGTAT  
 181 TTTCAGGAC ATATGAGATC TGTCGACTGG ATCCGGTAC GAATTGCTT ACTAAAAGCC  
 241 AGATAAACAGT ATGCGTATTT GCGCGCTGAT TTTGCGGTAA TAAGAATATA TACTGATATG  
 301 TATACCCGAA GTATGTCAAA AAGAGGTGTG CTTCTAGAAT GCAGTTAAG GTTACACCT  
 361 ATAAAAGAGA GAGCCGTTAT CGTCTGTTTG TGGATGTACA GAGTGTATT ATTGACACGC  
 421 CCGGGCGACG GATAGTGATC CCCCTGGCCA GTGCACGCTC GCTGTCAGAT AAAGTCTCCC  
 481 GTGAACTTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA  
 541 TGGCCAGTGT GCGGCTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA  
 601 ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGAAT ATAGAATTG CGGGCGCACT  
 661 CGAGATATCT AGACCCAGCT TTCTGTACA AGTTGGCAT TATAAGAAAG CATTGCTTAT  
 721 CAATTTGTTG CAACGAACAG GTCACATATCA GTCAAAATAA AATCATTATT TGCCATCCAG  
 781 CTGCAGCTCT GGCCCGTGTCA TCACAACTTC TGATGTTACA TTGCACAAAGA TAAAAATATA  
 841 TCATCATGAA CAATAAAACT GTCTGTTAC ATAAACAGTA ATACAAGGGG TGTATGAGC  
 901 CATATTCAAC GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT  
 961 GGGTATAAAAT GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGTAT  
 1021 GGGAAAGCCCG ATGCGCCAGA GTTGTTCCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT  
 1081 GTTACAGATG AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCTCT TCCGACCATC  
 1141 AAGCATTTTA TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCAGAAAA  
 1201 ACAGCATTC AGGTATTAGA AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG  
 1261 GCAGTGTCCC TGCGCCGGTT GCATTCGATT CCTGTTGTAA ATTGTCTTT TAACAGCGAT  
 1321 CGCGTATTTG GTCTCGCTCA GGCGCAATCA CGAATGAATA ACGGTTGGT TGATGCGAGT  
 1381 GATTTGATG ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAA  
 1441 CTTTTGCCAT TCTCACCGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT  
 1501 ATTTTTGAGC AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC  
 1561 CGATACCAGG ATCTTGCCAT CCTATGGAAC TGCCCTCGGT AGTTTCTCC TTCATTACAG  
 1621 AAACGGCTTT TTCAAAAATA TGGTATTGAT AATCTGATA TGAATAAAATT GCAGTTTCAT  
 1681 TTGATGCTCG ATGAGTTTTT CTAATCAGAA TTGGTTAATT GTGTTAACAA TTATTCAGAT  
 1741 TGGGCCCCCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT  
 1801 CTTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCAACCGCT ACCAGCGGTG  
 1861 GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTGG CTTCAAGCAGA  
 1921 GCGCAGATAAC CAAATACTGT TCTTCTAGTG TAGCGTAGT TAGGCCACCA CTTCAAGAAC  
 1981 TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCACTGGC TGCTGCCAGT  
 2041 GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG  
 2101 CGGTGGGCT GAACGGGGGG TTCTGTCACA CAGCCCCAGCT TGGAGCGAAC GACCTACACC  
 2161 GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCCCA CGCTTCCCGA AGGGAGAAAG  
 2221 CGGGACAGGT ATCCGGTAAG CGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA  
 2281 GGGGGAAACG CCTGGTATCT TTATAGCTCT GTCCGGTTTC GCCACCTCTG ACTTGAGCGT  
 2341 CGATTTTTGT GATGCTCGTC AGGGGGGGCG AGCCTATGGA AAAACGCCAG CAACCGGGCC  
 2401 TTTTACGGT TCCTGGGCTT TTGCTGGCCT TTTGCTACCA TGTTCTTCC TGCGTTATCC  
 2461 CCTGATTCTG TGGATAACCG TATTACCGCT AGCATGGATC TCAGGGACGT CTAACTACTA  
 2521 AGCGAGAGTA GGGAACTGCC AGGCATCAA TAAAACGAAA GGCTCAGTCG GAAGACTGGG  
 2581 CTTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCTG GAGTAGGACA AATCCGCCGG  
 2641 GAGCGGATTG GAACGTGTC AGCAACGGC CGGGAGGGTG GCGGGCAGGA CGCCCGCCAT  
 2701 AACTGCCAG GCATCAAACG AAGCAGAAGG CCATC

FIGURE 18B

28/240

Figure 19A: Cloning sites of the Entry Vector pENTR10

Int attL1 S.D. - 12 NdeI

--- tgg tac aaa aaa gca ggc ttc gaa cta agg aaa tac tta cat  
 --- dgg atg ttt tcc cgt ccg ang ett gat tcc ttt atg aat gta  
 Leu Tyr Lys Lys Ala Gly Phe Glu Leu Arg Lys Tyr Leu His

Kpn Xba Sph Bam Kpn EcoRI

atg gga acc aat tca ggc gac tgg atc egg tac cga att cgc ---  
 tcc cct tgg tta agt cag ctt acc tag gcp atg get taa geg ---  
 Met Gly Thr Asn Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI Not Xba EcoRI Xba att 2

Death --- aca att cgc ggc cgc act cga gat atc tag acc cag  
 (ccdB) --- tct taa gcg ccc gcg tga gct cta tag atc tgg gtc

Int

--- tcc tgg aca aca ---  
 gaa aga aca rot ruc ---

29/240

## pENTR10 2738 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAATTGAT  
 121 AAGCAATGCT TTTTATAAT GCCAATTG TACAAAAAAAG CAGGCTTCGA ACTAAGGAAA  
 181 TACTTACATA TGGGAACCAA TTCAGTCGAC TGGATCCGGT ACCGAATTCG CTTACTAAAA  
 241 GCCAGATAAAC AGTATGCGTA TTTGCGCGCT GATTTTGCCTG GTATAAGAAT ATATACTGAT  
 301 ATGTATACCC GAAGTATGTC AAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA  
 361 CCTATAAAAAG AGAGAGCCGT TATCGCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA  
 421 CGCCCCGGGGCG ACGGATGGTG ATCCCCCTGG CCAGTGCACCG TCTGCTGTCA GATAAAGTCT  
 481 CCCGTGAACCT TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG  
 541 ATATGGCCAG TGTGCCGGTC TCCGTTATCG GGAAAGAAGT GGCTGATCTC AGCCACCGCG  
 601 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTTCTGGGG AATATAGAAT TCAGGGCCGC  
 661 ACTCGAGATA TCTAGACCA GCTTTCTTGT ACAAAAGTTGG CATTATAAGA AAGCATTGCT  
 721 TATCAATTG TTGCAACGAA CAGGTCACTA TCAGTCAAAAA TAAAATCATT ATTTGCCATC  
 781 CAGCTGCAGC TCTGGCCCGT GTCTCAAAAT CTCTGATGTT ACATTGACA AGATAAAAAT  
 841 ATATCATCAT GAACAATAAA ACTGTCTGCT TACATAAAACA GTAATACAAG GGGTGTATG  
 901 AGCCATATTC AACGGGAAAC GTGAGGCCG CGATTTAAATT CCAACATGGA TGCTGATTTA  
 961 TATGGGTATA AATGGGCTCG CGATAATGTC GGGCAATCAG GTGCGACAAAT CTATCGCTTG  
 1021 TATGGGAAGC CCGATGGCC AGAGTTGTTT CTGAAACATG GCAAAGGTTAG CGTTGCCAAT  
 1081 GATGTTACAG ATGAGATGGT CAGACTAAAC TGGCTGACGG AATTTATGCC TCTTCCGACC  
 1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGGA  
 1201 AAAACAGCAT TCCAGGTATT AGAAGAATAT CCTGATTCA GTGAAAATAT TGTTGATGCG  
 1261 CTGGCAGTGT TCCCTGCCCG GTTGCAATTG ATTCCCTGTTT GAAATTGTCCT TTTAACAGC  
 1321 GATCGCGTAT TTCGTCTCGC TCAGCGCAA TCACGAATGA ATAACGGTTT GGTTGATGCG  
 1381 AGTGATTTG ATGACGAGCG TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT  
 1441 AAACCTTGC CATTCTCACCG GGATTCACTG GTCACTCATG GTGATTTCTC ACTTGATAAC  
 1501 CTTATTTTG ACGAGGGGAA ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA  
 1561 GACCGATACC AGGATCTGC CATCCATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA  
 1621 CAGAAACGGC TTTTCAAA ATATGGTATT GATAATCCTG ATATGAATAA ATTGAGTTT  
 1681 CATTGATGC TCGATGAGTT TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA  
 1741 GATTGGGCCCG CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA  
 1801 GATCCTTTT TTCTGCCCGT AATCTGCTGC TTGCAAACAA AAAAACCAAC GCTACCAGCG  
 1861 GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC  
 1921 AGAGCGCAGA TACCAAATAC TGTCTCTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG  
 1981 AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC  
 2041 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG  
 2101 CAGCGGTCTGG GCTGAACGGG GGGTCTGTC ACACAGCCCA GCTTGGAGCG AACGACCTAC  
 2161 ACCGAACCTGA GATACTACCA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA  
 2221 AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCCAC GAGGGAGCTT  
 2281 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CTCGCTGGGGT TTGCCCACCT CTGACTTGAG  
 2341 CGTCGATTTT TGTGATGTC GTCAGGGGGGG CGGAGCCSTAT GGAAAAAACGC CAGCAACCGC  
 2401 GCCTTTTAC GGTTCTGGC TCTTGTCTGG CCTTTTGCTC ACATGTTCTT TCCTGGTTA  
 2461 TCCCCTGATT CTGTGGATAA CCGTATTACC GCTAGCATGG ATCTCGGGGA CGTCTAACTA  
 2521 CTAAGCGAGA GTAGGGAACCT GCCAGGCATC GAATAAAAAG AAAGGCTCAG TCGGAAGACT  
 2581 GGGCCTTCG TTTTATCTGT TGTTTGTCTGG TGAACGCTCT CCTGAGTAGG ACAAAATCCGC  
 2641 CGGGAGCGGA TTTGAACGTT GTGAAGCAAC GGCCCCGGAGG GTGGCGGGCA GGACGCCCCGC  
 2701 CATAAAACTGC CAGGCATCAA ACTAAGCAGA AGGCCATC

FIGURE 19B

30/240

**Figure 20A: Cloning Sites of the Entry Vector pENTR11**

Int attL1 S.D. Kozak XmnI S.D.

TTG TAC AAA AAA GCA GGC TTC GAA GGA GAT AGA ACC AAT TCT CTA AGG AAA TAC  
 AAC ATG TTT TTT CGT CCG AAG CTT CCT CTA TCT TGG TTA AGA GAT TCC TTT ATG

Leu Tyr Lys Lys Ala Gly Phe Glu Gly Asp Arg Thr Asn Ser Leu Arg Lys Tyr

Kozak NcoI SalI BamHI KpnI EcoRI EcoRI NotI

TTA ACC ATG CTC GAC TGG ATC CGG TAC CGA ATT C-- ccdB --G AAT TCG CGG CCG  
 AAT TGG TAC CAG CTG ACC TAG GCC ATG GCT TAA G C TTA AGC GCC GGC

Leu Thr Met Val Asp Trp Ile Arg Tyr Arg Ile Asn Ser Arg Pro

XbaI EcoRV XbaI Int attL2

CAC TCG AGA TAT CTA GAC CCA GCT TTC TTG TAC AAA G  
 GTG AGC TCT ATA GAT CTG GGT CGA AAG AAC ATG TTT C

His Ser Arg Tyr Leu Asp Pro Ala Phe Leu Tyr Lys

31/240

## pENTR11 2744 bp (rotated to position 2578)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
348..653	ccdB
683..781	attL2
904..1713	KmR
1818..2391	ori

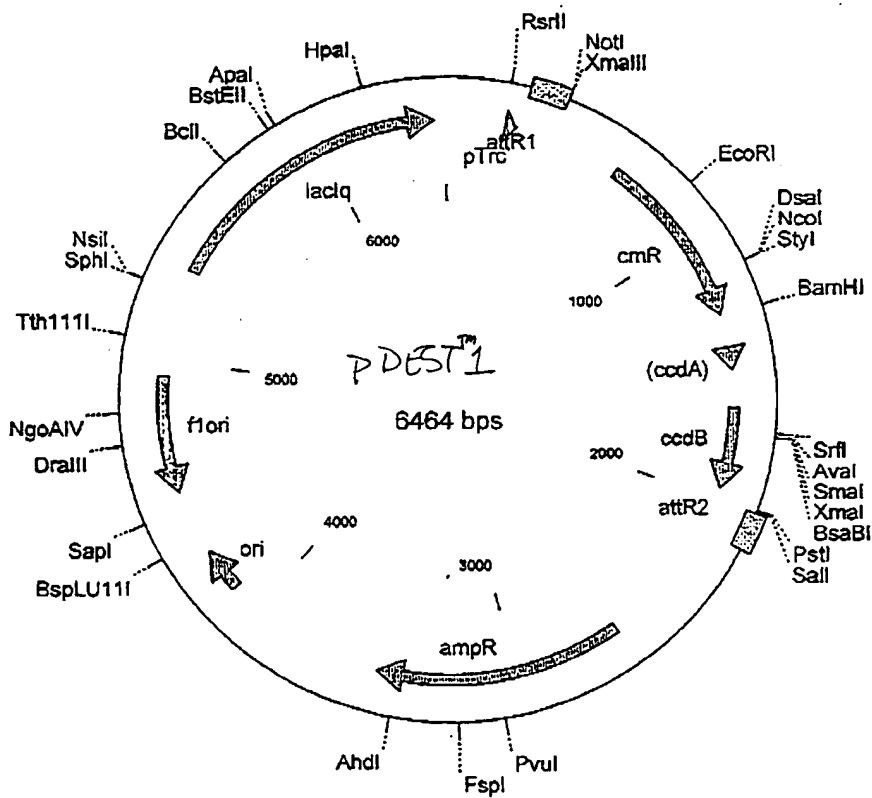
1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGITAGTT ACTTAAGCTC  
 61 GGGCCCCAAA TAATGATTTC ATTTGACTG ATAGTGCACCT GTTCGTTGCA ACAAATTGAT  
 121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAAG CAGGCTTCGA AGGAGATAGA  
 181 ACCAATTCTC TAAGGAAATA CTTAACCATG GTCGACTGGA TCCGGTACCG AATTGCTTA  
 241 CTAAAAGCCA GATAACAGTA TGCGTATTTG CGCGCTGATT TTTGCGGTAT AAGAATATAT  
 301 ACTGATATGT ATACCCGAAG TATGTCAAAAA AGAGGTGTGC TTCTAGAATG CAGTTTAAGG  
 361 TTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTG GGATGTACAG AGTGATATTA  
 421 TTGACACGCC CGGGCAGCG ATAGTGATCC CCCTGGCCAG TGCACGTCTG CTGTCAGATA  
 481 AAAGTCTCCCG TGAACTTAC CCGGGGGTGC ATATCGGGGA TGAAAGCTGG CGCATGATGA  
 541 CCACCGATAT GGCCAGTGTG CGGGCTCCG TTATCGGGGA AGAAGTGGCT GATCTCAGCC  
 601 ACCCGGAAAAA TGACATCAAA AACGCCATTA ACCTGATGTT CTGGGGAAATA TAGAATTGCG  
 661 GCGCCGACTC GAGATATCTA GACCCAGCTT TCTTGTACAA AGTTGGCATT ATAAGAAAGC  
 721 ATTGCTTATC AATTGTTGCA AACGAACAGG TCACTATCAC TCAAAATAAA ATCAATTATTT  
 781 GCCATCCAGC TGCAGCTCTG GCCCCTGTCT CAAAATCTCT GATGTTACAT TGCAACAAGAT  
 841 AAAAATATAT CATCATGAAC AATAAAACTG TCTGCTTACA TAAACAGTAA TACAAGGGGT  
 901 GTTATGAGCC ATATTCAACG GGAAACGTG AGGCGCGCAT TAAATTCCAA CATGGATGCT  
 961 GATTTATATG GGTATAAATG GGCTCGCGAT ATATGCGGGC AATCAGGTGC GACAATCTAT  
 1021 CGCTTGTATG GGAAGCCCGA TGCGCCAGAG TTGTTCTGA AACATGGCAA AGGTAGCGTT  
 1081 GCCAATGATG TTACAGATGA GATGGTCAGA CTAAACTGGC TGACGGAATT TATGCCCTTT  
 1141 CCGACCATCA AGCATTCTAT CGCTACTCCT GATGATGCGAT GGTTACTCAC CACTGCGATC  
 1201 CCCGGAAAAA CAGCATTCCA GGTATTAGAA GAATATCCTG ATTCAGGTGA AAATATTGTT  
 1261 GATGCGCTGG CAGTGTCTC GCGCCGGTTG CATTGATTC CTGTTTGAA TTGCTCTTTT  
 1321 AACAGCGATC GCGTATTTG TCTCGCTCAG GCGCAATCAC GAATGAATAA CGGTTGGTT  
 1381 GATGCGAGTG ATTTGATGA CGAGCGTAAT GGCTGGCCTG TTGAACAAGT CTGAAAGAA  
 1441 ATGCATAAAC TTTTGCCATT CTCACCGGAT TCAGTCGTCA CTCATGGTGA TTTCTCACTT  
 1501 GATAACCTA TTTTGACGA GGGGAAATTAA ATAGGTTGTA TTGATGTTGG ACGAGTCGGA  
 1561 ATCGCAGACC GATACCAGGA TCTTGCCATC CTATGGAACG GCCTCGGTGA GTTTCTCCT  
 1621 TCATTACAGA AACGGCTTT TCAAAATAT GGTATTGATA ATCCTGATAT GAATAATTG  
 1681 CAGTTTCATT TGATGCTCGA TGAGTTTTC TAATCAGAAT TGGTTAATTG GTGTAACAT  
 1741 TATTGAGATT GGGCCCCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT  
 1801 TCTTGAGATC CTTTTTTCTC GCGCGTAATC TGCTGTTGC AAACAAAAAA ACCACCGCTA  
 1861 CCAGCGGTGG TTGTTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAACGTGGC  
 1921 TTCAGCAGAG CGCAGATAACC AAATACTGTT CTTCTAGTGT AGCCGTAGTT AGGCCACCAC  
 1981 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT  
 2041 GCTGCCAGTG GCGATAAGTC GTGCTTACCG GGGTTGGACT CAAGACGATA GTTACCGGAT  
 2101 AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG  
 2161 ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCGAA  
 2221 GGGAGAAAGG CGGACAGGT A TCCGCTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG  
 2281 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATGTCCTG TCGGGTTTCG CCACCTCTGA  
 2341 CTTGAGCGTC GATTTTGTC ATGTCGTCA GGGGGCGGA GCCTATGGAA AAACGCCAGC  
 2401 AACCGGGCCT TTTTACGGTT CCTGGCCTT TGCTGGCCTT TTGCTCACAT GTTCTTCCCT  
 2461 GCGTTATCCC CTGATTCTGT GGATAACCCT ATTACCGCTA GCATGGATCT CGGGGACGTC  
 2521 TAACTACTAA GCGAGAGTAG GGAACCTGCA GGCATCAAAT AAAACGAAAG GCTCAGTCGG  
 2581 AAGACTGGGC CTTTCGTTT ATCTGTTGTT TGTCGGTGAA CGCTCTCCTG AGTAGGACAA  
 2641 ATCCGCCGGG AGCGGATTIG AACGTTGTA AGCAACGGCC CGGAGGGTGG CGGGCAGGAC  
 2701 GCCCGCCATA AACTGCCAGG CATCAAACTA AGCAGAAGGC CATC

FIGURE 20B

32 | 240

## **Figure : z/1A:pDEST1 Native Protein Expression in E. coli**

-35 Trc promoter -10 mRNA  
 1 atgagctgt gacaattaat catccggctc gtataatgtg tggatttgt agccggataac  
 tactcgacaa ctgttaatta gtaggccgag catattacac accttaaacac tcgcctattg  
 61 aatttcacac agaaaacaga caggatagg atcacaagtt tgatgttata agctgttacgaa  
 ttaaagtgtg tcctttgtct gtcacatatcc taggttcaaa acatgtttttt ttgttttgtt



33/240

## pDEST1 6464 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
216..257	Trc promoter
397..273	attR1
647..1306	CmR
1426..1510	inactivated ccdA
1648..1953	ccdB
1994..2118	attR2
2598..3503	ampR
4104..4264	ori
4504..4941	flori (f1 intergenic region)
5340..6420	lacIq

1 GTTTGACAGC TTATCATCGA CTGCACGGTG CACCAATGCT TCTGGCGTCA GGCAGGCCATC  
 61 GGAAGCTGTG GTATGGCTGT GCAGGTCGA AATCACTGCA TAATTCTGT CGCTCAAGGC  
 121 GCACTCCCGT TCTGGATAAT GTTTTTGCG CCGACATCAT AACGGTTCTG GCAAATATT  
 181 TGAAATGAGC TGTTGACAAT TAATCATTCCG GTCCGTATAA TCTGTGGAAT TGTGAGCGGG  
 241 ATAACAATTT CATCGCAGG TACCAAGCTA TCACAAGTT GTACAAAAAA GCTGAACGAG  
 301 AAACGTAAAA TGATATAAAT ATCAATATAT TAAATTAGAT TTGCTATAA AAACAGACTA  
 361 CATAATACTG TAAAACACAA CATATCCAGT CACTATGGCG GCCGCTAACG TGGCAGCATT  
 421 ACCCGACGCA CTTTGCGCCG AATAAAATACC TGTGACGGAA GATCACTTCG CAGAATAAAT  
 481 AAATCCTGGT GTCCCTGTTG ATACCGGGAA GCCCTGGGCC AACTTTGGC GAAAATGAGA  
 541 CGTTGATCGG CACGTAAGAG GTTCCAACCT TCACCATAAT GAAAATAAGAT CACTACCGGG  
 601 CGTATTTTTG GAGTTATCGA GATTTTCAGG AGCTAAGGAA GCTAAAATGG AGAAAAAAAT  
 661 CACTGGATAT ACCACCGTTG ATATATCCC ATGGCATTG AAAGAACATT TTGAGGCATT  
 721 TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCGTTAG CTGGATATTAA CGGCCTTTT  
 781 AAAGACCGTA AAGAAAAATA AGCACAAGTT TTATCCGGCC TTATTACACA TTCTTGCCCCG  
 841 CCTGATGAAT GCTCATCCCG AATTCCGTAT GGCAATGAAA GACGGTGAGC TGGTGATATG  
 901 GGATAGTGTG CACCCCTGTT ACACCGTTT CCATGAGCAA ACTGAAACGT TTTCATCGCT  
 961 CTGGAGTGA TACCACGACG ATTTCGGCA GTTCTACAC ATATATTGCG AAGATGTGGC  
 1021 GTGTTACGGT GAAAACCTGG CCTATTCCC TAAAGGGTTT ATTGAGAATA TGTTTTCGT  
 1081 CTCAGCCAAT CCCTGGGTGA GTTTCACCAAG TTTGATTTA AACGTGGCCA ATATGGACAA  
 1141 CTTCTTCGCC CCCGTTTCA CCATGGCAA ATATTATACG CAAGGCGACA AGGTGCTGAT  
 1201 GCGCCTGGCG ATTCAAGTTC ATCATGCCGT CTGTGATGGC TTCCATGTCG GCAGAATGCT  
 1261 TAATGAATTA CAACAGTACT GCGATGAGTG GCAGGGCGGG GCGTAAACGC GTGGATCCGG  
 1321 CTTACTAAAA GCCAGATAAC AGTATGGCTA TTGCGCGCT GATTTTGCG GTATAAGAAT  
 1381 ATATACTGAT ATGTATACCC GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT  
 1441 ACAGTACAG TTGACAGCGA CAGCTATCAG TTGCTCAAGG CATATATGAT GTCAATATCT  
 1501 CCGGTCTGGT AAGCACAACC ATGCAGAATG AAGCCCGTCG TCTGCGTGC GAACGCTGGA  
 1561 AAGCGGAAAA TCAGGAAGGG ATGGCTGAGG TCGCCCGGTT TATTGAAATG AACGGCTCTT  
 1621 TTGCTGACGA GAACAGGGAC TGGTGAATG CAGTTAAGG TTTACACCTA TAAAAGAGAG  
 1681 AGCCGTTATC GTCTGTTCTG GGATGTACAG AGTGTATTTA TTGACACGCC CGGGCGACGG  
 1741 ATGGTGATCC CCCTGGCCAG TGACACGTCG CTGTCAGATA AAGTCTCCG TGAACCTTAC  
 1801 CCGGTGGTGC ATATCGGGGA TGAAAGCTGG CGCATGATGA CCACCGATAT GGCCAGTGTG  
 1861 CCGGTCTCCG TTATCGGGGA AGAAGTGGCT GATCTCAGCC ACCCGAAAAA TGACATCAA  
 1921 AACGCCATTA ACCTGATGTT CTGGGAAATA TAAATGTCAG GCTCCCTTAT ACACAGCCAG  
 1981 TCTGCAGGTC GACCATAGTG ACTGGATATG TTGTGTTTA CAGTATTATG TAGTCTGTTT  
 2041 TTTATGAAA ATCTAATTTA ATATATTGAT ATTATATC A TTTACGTTT CTCGTTCA  
 2101 TTTCTTGTAC AAAGTGGTGA TAGCTTGGCT GTTTGGCGG ATGAGAGAAAG ATTTTCA  
 2161 TGATACAGAT TAAATCAGAA CGCAGAACGG GTCTGATAAA ACAGAATTG CCTGGCGGCA  
 2221 GTAGCGCGGT GGTCCCACCT GACCCCATGC CGAACCTCAGA AGTGAACACGC CGTAGCGCCG  
 2281 ATGGTAGTGT GGGGTCTCCC CATCGCAGAG TAGGAAACTG CCAGGCATCA AATAAACGA  
 2341 AAGGCTCAGT CGAAAGACTG GGCCTTTCGT TTTATCTGTT GTTGTGCGGT GAACGCTCTC  
 2401 CTGAGTAGGA CAAATCCGCC GGGAGCGGGAT TTGAACGTT CGAAGCAACG GCCCGGAGGG  
 2461 TGGCGGGCAG GACGCCGCC ATAAACTGCC AGGCATCAA TAAAGCAGAA GGCCATCCTG  
 2521 ACGGATGGCC TTTTGCCTT TCTACAAACT CTTTTGTTT ATTTCCTAA ATACATTCAA-

FIGURE 21B

34/240

2581 ATATGTATCC GCTCATGAGA CAATAACCT GATAAAATGCT TCAATAATAT TGAAAAAGGA  
 2641 AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTTGCC GCATTTGCC  
 2701 TTCCTGTTT TGCTCACCCA GAAACGCTGG TGAAAGTAA AGATGCTGAA GATCAGTTGG  
 2761 GTGCACCGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTC  
 2821 GCCCGAAGA AGTTTTCGA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT  
 2881 TATCCCGTGT TGACGCCGG CAAGAGCAAC TCGGTGCGC Catacaactat TCTCAGAACATG  
 2941 ACTGGTTGA GTCATCACA GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG  
 3001 AATTATGCAG TGTGCAATA ACCATGAGTG ATAACACTGC GGCCAACCTA CTTCTGACAA  
 3061 CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTCGCACAA CATGGGGGAT CATGTAACCTC  
 3121 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA  
 3181 CGATGCCTAC AGCAATGGCA ACAACGTTGC GCAAACATTAA AACTGGCGAA CTACTTAACCTC  
 3241 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGAGGGCGGA TAAAGTTGCA GGACCACTTC  
 3301 TGCCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG  
 3361 GGTCTCGCG TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCCTCCGT ATCGTAGTTA  
 3421 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG  
 3481 GTGCCCTACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA  
 3541 TTGATTTAAA ACTTCATTAA TAATTTAAA GGATCTAGGT GAAGATCCTT TTGATAATC  
 3601 TCATGACCAA AATCCCTAA CGTGAGTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA  
 3661 AGATCAAAGG ATCTTCTTGA GATCCTTTT TTCTGCGCGT AATCTGCTG TTGCAAACAA  
 3721 AAAAACACC GCTACCAGCG GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTC  
 3781 CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCTTCTA GTGTAGCCGT  
 3841 AGTTAGGCCA CCACTTCAG AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC  
 3901 TGTACACGAG GGCTGCTGCC AGTGGCGATA AGTCGTGTC TACCGGGTTG GACTCAAGAC  
 3961 GATAGTTACC GGATAAGGGCG CAGCGGTGG GCTGAACGGG GGGTCTGTC ACACAGCCCA  
 4021 GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG  
 4081 CCACGCTTCC CGAAGGGGAGA AAGGGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG  
 4141 GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACCGCTGGTA TCTTTATAGT CCTGTCGGGT  
 4201 TTCGCCACCT CTGACTTGAG CGTCGATTT TGTGATGTC GTCAGGGGG CGGAGCCTAT  
 4261 GGAAAAACGC CAGCAACCGC GCCTTTTAC GGTCTCTGGC TTITGCTGG CCTTTTGCTC  
 4321 ACATGTTCTT TCCCTGCTTA TCCCTGATT CTGTTGATAA CCGTATTACCC GCCTTTGAGT  
 4381 GAGCTGATAC CGCTCGCCG AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG  
 4441 CGGAAGAGCG CCTGATGCGG TATTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA  
 4501 TAATTTGTT AAAATTGCGG TTAAATTTTT GTTAAATCAG CTCATTTTT AACCAATAGG  
 4561 CGAAATCGG CAAAATCCCT TATAAATCAA AAGAATAGAC CGAGATAGGG TTGAGTGTG  
 4621 TTCCAGTTTG GAACAAGAGT CCACTATTAA AGAACGTGGA CTCCAACGTC AAAGGGCGAA  
 4681 AAACCGTCTA TCAGGGCGAT GGCCCACAC TGGAACCATC ACCCTAATCA AGTTTTTGG  
 4741 GGTGAGGTG CCGTAAAGCA CTAAATCGGA ACCCTAAAGG GAGCCCCGA TTTAGAGCTT  
 4801 GACGGGGAAA GCGGGCGAAC GTGGCGAGAA AGGAAGGGAA GAAAGCGAAA GGAGCGGGCG  
 4861 CTAGGGCGCT GGCAAGTGTG GCGGTCAAC TGCGCGTAAC CACCACACCC GCCGCGCTTA  
 4921 ATGCGCCGCT ACAGGGCGCG TCCATTGCG ATTCAAGGCTG CTATGGTCA CTCTCAGTAC  
 4981 AATCTGCTCT GATGCCGCAT AGTTAAGCCA GTACCGAGTC CGTAGCGATA TCGGAGTGTG  
 5041 TACACTCCG TATCGCTACG TGACTGGTC ATGGCTCGC CCGACACCC GCCAACACCC  
 5101 GCTGACGCGC CCTGACGGGC TTGTCGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC  
 5161 GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGCTAT CACCGAAACG CGCGAGGCAG  
 5221 CAGATCAATT CGCGCGCGA GGCAGCGGG CATGCATTAA CGTGTGACACC ATCGAATGGT  
 5281 GCACAAACCTT TCCGGTATG GCATGATAGC GCGCGGAAGA GAGTCAAATC AGGGTGTG  
 5341 ATGTGAAACC AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCC  
 5401 TTTCCCGCGT GGTAACCGAG GCCAGCGACG TTCTGCGAA AACCGGGAA AAAGTGGAAAG  
 5461 CGCGCATGGC GGAGCTGAAT TACATCCCA ACCCGCTGGC ACAACAACTG GCGGGCAAC  
 5521 AGTCGTTGCT GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACGGCCG TCGCAAATTG  
 5581 TCGCGCGAT TAAATCTCGC GCGCATCAAC TGCGGAGCA CGTGGTGGTG TCGATGGTAG  
 5641 AACGAAGCGG CGTCAAGCC TGTAAGCGG CGGTGACAA TCTTCTCGCG CAACCGCTCA  
 5701 GTGGGCTGAT CATTAACTAT CGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT  
 5761 GCACTAATGT TCCGGCGTTA TTCTTGTGAT TCTCTGACCA GACACCCATC AACAGTATTG  
 5821 TTTCTCCCA TGAAGACGGT ACGCGACTGG CGGTGGAGCA TCTGGTCGA TTGGGTCA  
 5881 AGCAAATCGC GCTGTTAGCG GGCCCAATTAA GTTCTGTC CCGCGCTCT CGTCTGGCTG  
 5941 GCTGGCATAA ATATCTCACT CGCAATCAAA TTCAGCCGAT AGCGGAACGG GAAGGGCACT  
 6001 GGAGTGCCT GTCCGGTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCCA-

FIGURE 21C

35/240

6061 CTGCGATGCT GGTTGCCAAC GATCAGATGG CGCTGGGCAC AATGCGGCC ATTACCGAGT  
6121 CCGGGCTGCG CGTTGGTGC GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT  
6181 CATGTTATAT CCCGCCGTTA ACCACCATCA AACAGGATT TCGCCTGCTG GGGCAAACCA  
6241 GCGTGGACCG CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC  
6301 CCGTCTCACT GGTGAAAAGA AAAACACCC TGCGACCCAA TACGAAACCC GCCTCTCCCC  
6361 GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC  
6421 AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CGCGAATTGA TCTG

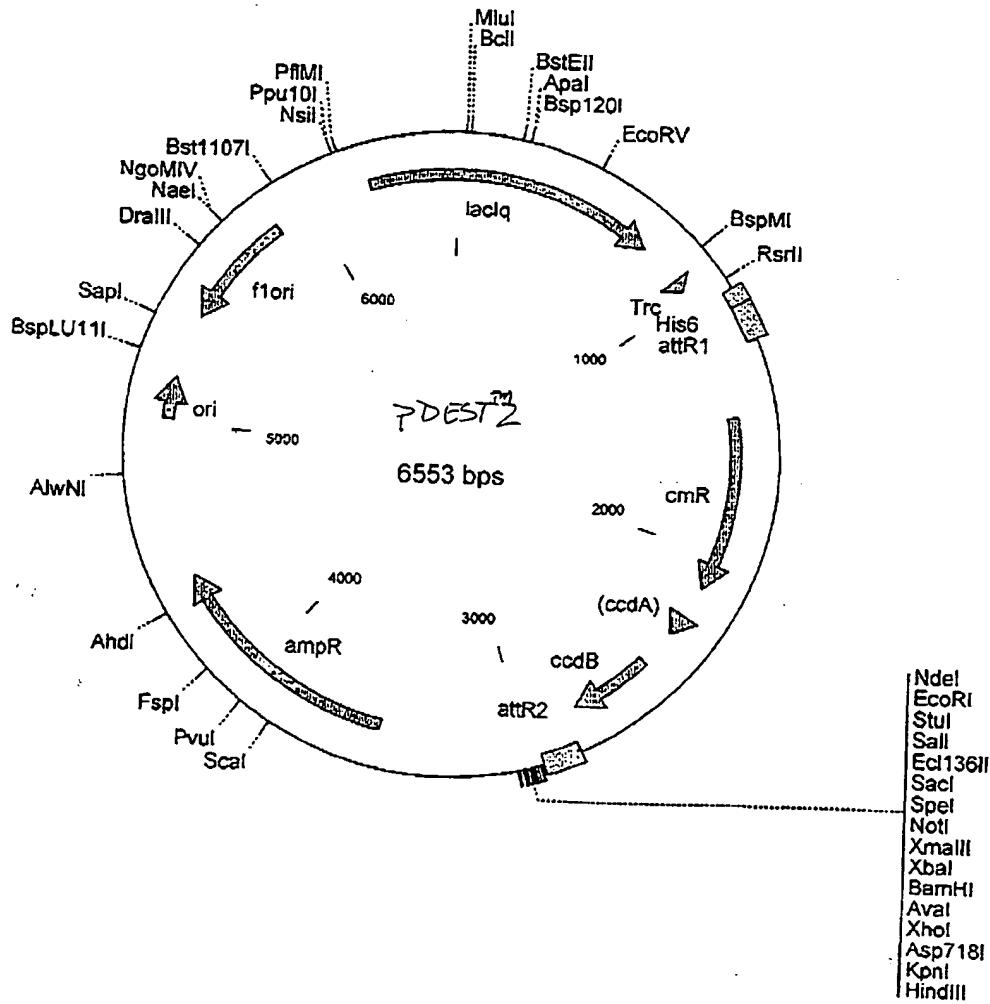
FIGURE 21D

36/240

**Figure 22A:** pDEST2

## His6 fusions in E. coli

970   aat att ctg aaa tga gct gtt gac aat tad tca tcc ggt ccg atc aat ctg  
 tta taa gac ttt act cga caa ctg tta att agt agg cca ggc ata tta gac  
 1021   tgg aat tgt gag cgg ata aca att tca cac agg aaa cag acc atg tcg tdc  
 acc tta aca ctc gcc tat tgt taa agt gtg tcc ttt gtc tgg tac agc atg  
 1072   Tyr His His His His His Gly Tie Tyr Ser Int ATP R1  
 tac cat cac cat ctc cat cat ggt att acc agt tgg taa aaa gct oaa  
 atg gta gtg gta gtg ccg tag tgt tca aac atg ttt ctt csa cxt



37/260

## pDEST2 6553 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
912..962	Trc
1223..1009	attR1
1473..2132	CmR
2252..2336	inactivated ccdA
2474..2779	ccdB
2820..2944	attR2
3509..4414	ampR
5015..5175	ori
5415..5852	flori (f1 intergenic region)
6225..752	lacIq

1 GGCGGTGCAC AATCTTCTCG CGCAACGCGT CAGTGGGCTG ATCATTAACT ATCCGCTGGA  
 61 TGACCAGGAT GCCATTGCTG TGGAAAGCTGC CTGCACTAAT GTTCCGGCGT TATTCTTGA  
 121 TGCTCTGAC CAGACACCCA TCAACAGTAT TATTTCTCC CATGAAGACG GTACCGACT  
 181 GGGCGTGGAG CATCTGGTCG CATTGGTCA CCAGCAAATC GCGCTGTTAG CGGGCCATT  
 241 AAGTTCTGTC TC GGCGCGTC TGCGTCTGGC TGGCTGGCAT AAATATCTCA CTCGAATCA  
 301 AATTCAAGCCG ATAGCGGAAC GGGAAAGCGA CTGGAGTGC ATGTCGGTT TTCAACAAAC  
 361 CATGCAAATG CTGAATGAGG CAATCGTTCC CACTCGATG CTGGTTGCCA ACGATCAGAT  
 421 GGGCGTGGGC GCAATGCGCG CCATTACCGA GTCCGGGCTG CGCGTTGGTG CGGATATCTC  
 481 GGTAGTGGGA TACGACGATA CCGAAGACAG CTCATGTTAT ATCCCAGCGT CAACCACCAT  
 541 CAAACAGGAT TTTCGCCTGC TGGGGCAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA  
 601 GGGCCAGGCG GTGAAGGGCA ATCAGCTGTT GCCGCTCTCA CTGGTAAAAA GAAAACAC  
 661 CCTGGCACCC AATACGCAA CCGCCTCTCC CGCGCGTTG GCGGATTCAAT TAATGCAGCT  
 721 GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGGAGCG CAACGCAATT AATGTGAGTT  
 781 AGCGCGAATT GATCTGGTTT GACAGCTTAT CATGACTGC ACGGTGCACC AATGCTTCTG  
 841 GCGTCAGGCA GCCATCGGAA GCTGTGGTAT GGCTGTGCAG GTCGTAATC ACTGCATAAT  
 901 TCGTGTGCT CAAGGCGCAC TCCCGTTCTG GATAATGTTT TTGCGCCGA CATCATAACG  
 961 GTTCTGGCAA ATATTCTGAA ATGAGCTGTT GACAATTAAT CATCCGGTCC STATAATCTG  
 1021 TGGAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGA CCATGTCGTA CTACCATCAC  
 1081 CATCACCAC ACAGGCATCAC AAGTTGTAC AAAAAAGCTG AACGAGAAC GTAAAATGAT  
 1141 ATAAATATCA ATATATTAAA TTAGATTTG CATAAAAAAC AGACTACATA ATACTGTAAA  
 1201 ACACAACATA TCCAGTCACT ATGGCGGCG CTAAGTGGC AGCATCACCC GACGCACTTT  
 1261 GCGCCGAATA AATACCTGTG ACAGGAAGATC ACTTGCAGA ATAATAAAAT CCTGGTGTCC  
 1321 CTGTTGATAC CGGGAAAGCCC TGGGCAACT TTGGCGAAA ATGAGACGTT GATCGGCACG  
 1381 TAAGAGGTTT CAACTTTCAAC CATAATGAAA TAAGATCACT ACCGGGCGTA TTTTTGAGT  
 1441 TATCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT GGATATACCA  
 1501 CCGTTGATAT ATCCCAATGG CATCGAAAG AACATTTGA GGCATTTCACT TCAGTTGCTC  
 1561 AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTAAAG ACCGTAAGA  
 1621 AAAATAAGCA CAAGTTTAT CGGCCTTTA TTACATTCT TGCCCGCTG ATGAATGCTC  
 1681 ATCCGGAATT CCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC  
 1741 CTTGTTACAC CGTTTCCAT GAGCAAACCTG AACGTTTTC ATCGCTCTGG AGTGAATACC  
 1801 ACGACGATT CCAGCAGTTT CTACACATAT ATTGCAAGA TGTGGCGTGT TACGGTAAA  
 1861 ACCTGGCTA TTTCCCTAAA GGGTTATTG AGAATATGTT TTGCGCTCA GCGAACCCCT  
 1921 GGGTGAGTTT CACCAAGTTT GATTTAACG TGGCAATAT GGACAACCTC TTGCCCCCG  
 1981 TTTTCACCAT GGGCAAATAT TATACCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC  
 2041 AGGTTCATCA TGCGCTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT GAATTACAAC  
 2101 AGTACTGCGA TGAGTGGCAG GGGGGGCGT AAACGGCTGG ATCCGGCTTA CTAAAAGCCA  
 2161 GATAACAGTA TGCCTATTG CGCGCTGATT TTGCGGTAT AAGAATATAT ACTGATATGT  
 2221 ATACCCGAAG TATGTCAAAA AGAGGTGTGC TATGAAGCAG CGTATTACAG TGACAGTTGA  
 2281 CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA ATATCTCCGG TCTGGTAAGC  
 2341 ACAACCATGC AGAATGAAGC CGTCGCTCTG CGTGGCGAAC GCTGGAAAGC GGAAAATCAG  
 2401 GAAGGGATGG CTGAGGTGCG CCGGTTATT GAAATGAACG GCTCTTTGC TGACGAGAAC  
 2461 AGGGACTGGT GAAATGCAGT TTAAGGTTA CACCTATAAA AGAGAGAGCC GTTATCGTCT  
 2521 GTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCGGG CGACGGATGG TGATCCCCCT-

FIGURE 22B

38/240

2581 GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCCTGAA CTTTACCCGG TGGTGCATAT  
 2641 CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCGG TCTCCGTTAT  
 2701 CGGGGAAGAA GTGGCTGATC TCAGGCCACCG CGAAAATGAC ATCAAAAACG CCATTAACCT  
 2761 GATGTTCTGG GGAATATAAA TGTCAAGGCTC CCTTATACAC AGCCAGTCTG CAGGTCGACC  
 2821 ATAGTGAATG GATATGTTGT GTTTTACAGT ATTATGTAGT CTGTTTTTTA TGCAAAATCT  
 2881 AATTTAATAT ATTGATATTG ATATCATTT ACGTTTCTCG TTCAGCTTC TTGTACAAAG  
 2941 TGGTGTGATGCC CATATGGAA TTCAAAGGCC TAGCTCGACG AGCTCACTAG TCGCGGCCGC  
 3001 TTCTAGAGGA TCCCTCGAGG TAGCGGTAC CAAGCTTGGC TGTTTGGCG GATGAGAGAA  
 3061 GATTTTCAGC CTGATACAGA TTAATTCAGA ACGCAGAACG GGTCTGATAA AACAGAATTT  
 3121 GCCTGGCGGC AGTAGCGCGG TGGTCCACC TGACCCCATG CCGAACTCAG AAGTGAACAG  
 3181 CCGTAGCGCC GATGGTAGTG TGGGCTCTCC CCATGCGAGA GTAGGAAACT GCCAGGCATC  
 3241 AAATAAAACG AAAGGCTCAG TCGAAAGACT GGGCTTCTCG TTTTATCTGT TGTTTGTGG  
 3301 TGAACGCTCT CCTGAGTAGG ACAATCCGC CGGGAGCGGA TTGAACTGTT GCGAACAAAC  
 3361 GGCCCAGGAGG GTGGCGGGCA GGACGCCGC CATAAACTGC CAGGCATCAA ATTAAGCAGA  
 3421 AGGCCATCCT GACGGATGGC CTTTTGCGT TTCTACAAAC TCTTTTGT TATTTTCTA  
 3481 AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAAATGC TTCAATAATA  
 3541 TTGAAAAAAGG AAGAGTATGA GTATTCAACA TTTCCTGTC GCCCTTATTG CTTTTTTG  
 3601 GGCATTTGCTC TTTCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA  
 3661 AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT  
 3721 TGAGAGTTTT CGCCCCGAAG AACGTTTCC AATGATGAGC ACTTTAAAG TTCTGCTATG  
 3781 TGGCGCGGTA TTATCCCGT TTGACCCGGG GCAAGAGCAA CTGGTCGCC GCATACACTA  
 3841 TTCTCAGAAT GACTTGGTTG AGTACTCACC AGTCACAGAA AAGCATCTT CCGATGGCAT  
 3901 GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT  
 3961 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGACA ACATGGGGGA  
 4021 TCATGTAACG CGCCTTGATC GTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA  
 4081 GCGTGACACC ACGATGCCA CAGCAATGGC AACAACGTT CGAAACTAT TAATGGCGA  
 4141 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCCG ATAAAGTTGC  
 4201 AGGACCACCT CTGCGCTCGG CCCTTCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC  
 4261 CGGTGAGCGT GGGTCTCGG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG  
 4321 TATCGTAGTT ATCTACACCA CGGGAGTC GGCAACTATG GATGAACGAA ATAGACAGAT  
 4381 CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGTTAACCTG TCAGACCAAG TTTACTCATA  
 4441 TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAA AGGATCTAGG TGAAGATCCT  
 4501 TTTTGATAAT CTCATGACCA AAATCCCTA ACGTGAGTT TCCTTCAACT GAGCGTCAGA  
 4561 CCCCTAGAA AAGATCAAAG GATCTCTTG AGATCCTTT TTTCTGCGT TAATCTGCTG  
 4621 CTGCAAACA AAAAACCCAC CGCTACCAGC GGTGGTTTGT TTGCGGATC AAGAGCTACC  
 4681 AACTCTTTT CCGAAGGTTA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGCTTCT  
 4741 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACTCGC  
 4801 TCTGCTAATC CTGTTACAC TGGCTGTCAG CAGTGGCGAT AAGTCGTGTC TTACCGGGTT  
 4861 GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTG GGTGAACGG GGGGTTGCTG  
 4921 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT  
 4981 ATGAGAAAGC GCCACGCTTC CCGAAGGGG AAAGGCGGAC AGGTATCCGG TAAGGGCAG  
 5041 GGTGGAAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGA AACGGCTGGT ATCTTTATAG  
 5101 TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGCGATGCT CGTCAGGGGG  
 5161 GCGGAGCCTA TGGAAAAAACG CCAGCAACGC GGCCTTITA CGGTTCTGG CCTTTTGCTG  
 5221 GCCTTTGCT CACATGTTCT TTCTGCGTT ATCCCCGTAT TCTGTTGATA ACCGTATTAC  
 5281 CGCCTTGAG TGAGCTGATA CCGCTCGCCG CAGCGAACG ACCGAGGCCA GCGAGTCAGT  
 5341 GAGCGAGGAA GCGGAAGAGC GCGTGATGCG GTATTTCTC CTTACGCATC TGTGGGTAT  
 5401 TTACACCCGC ATAATTTTGT TAAAATTGCG GTTAAATTTT TTGTTAAATCA GCTCATT  
 5461 TAACCAATAG GCCGAAATCG GCAAAATCCC TTATAAAATCA AAAGAATAGA CCGAGATAGG  
 5521 GTTGTAGTGTGTT GTTCCAGTTT GGAACAAGAG TCCACTTATA AAGAACGTGG ACTCCAACGT  
 5581 CAAAGGGCGA AAAACCGTCT ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCCCTAAC  
 5641 AAGTTTTTG GGGTCGAGGT GCGTAAAGC ACTAAATCGG AACCTAAAG GGAGCCCCCG  
 5701 ATTTAGAGCT TGACGGGGAA AGCCGGCGAA CGTGGCGAGA AAGGAAGGGA AGAAAGCGAA  
 5761 AGGAGCGGGC GCTAGGGCCG TGGCAAGTGT AGCGGTACG CTGCGCTAA CCACCAACACC  
 5821 CGCCCGCTT AATGCGCCG TACAGGGCGC GTCCCATTG CCATTCAAGC TGCTATGGTG  
 5881 CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAACG CAGTATACAC TCCGCTATCG  
 5941 CTACGTGACT GGGTCATGGC TGCGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA  
 6001 CGGGCTTGTG TGCTCCCGGC ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC-

FIGURE 22C

39/240

6061 ATGTGTCAGA GGTTTCACC GTCATCACCG AAACGGCCGA GCCAGCAGAT CAATTCCGCG  
6121 GCGAAGGCAGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAAA ACCTTTCGCG  
6181 GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAAGGGT GGTGAATGTG AAACCAGTAA  
6241 CGTTATACGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTTCC CGCGTGGTGA  
6301 ACCAGGCCAG CCACGTTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGGCG ATGGCGGAGC  
6361 TGAATTACAT TCCCAACCGC GTGGCACAAAC AACTGGCGGG CAAACAGTCG TTGCTGATTG  
6421 GCGTTGCCAC CTCCAGTCTG GCCCTGCACG CGCCGTCGCA ATTGTCGCG GCGATTAAT  
6481 CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTGAT GGTAGAACGA AGCGGCGTCG  
6541 AAGCCTGTAA AGC

FIGURE 22D

40/240

Figure 23A: PDEST3

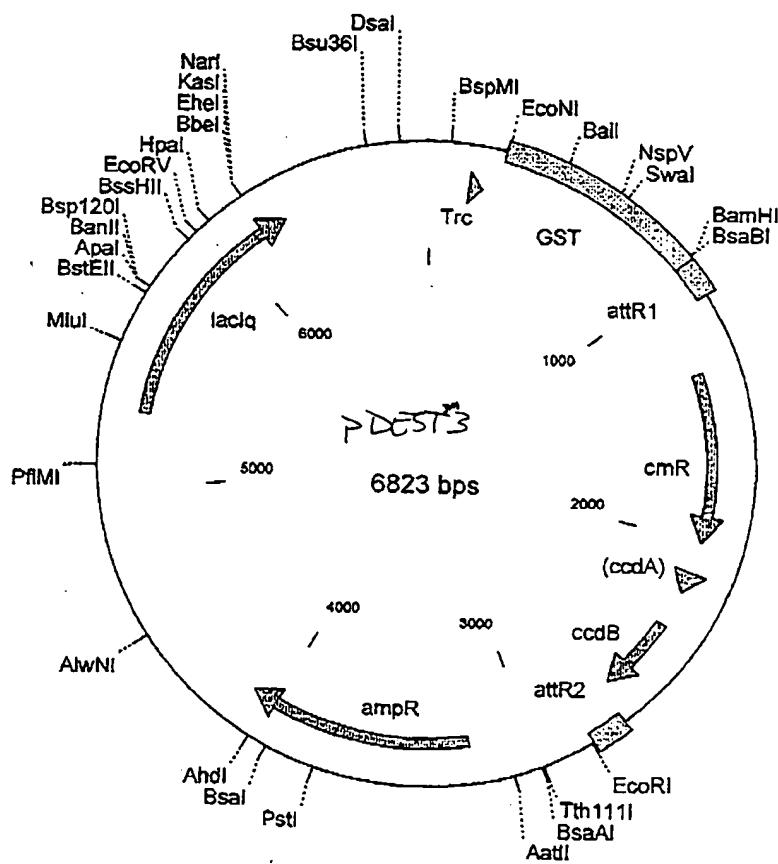
## GST fusions in E. coli

154 cggttc tgg caa ata ttc tga aat gag ctg -35 Trc promoter  
 gcc aag acc gtt tat aag act tta ctc gac aac tgt taa tta gta gcc gag  
 ~10 mRNA

205 gta caa gtgt gtg gaa ttg tga gcg gat aac aat ttc aca cag gaa aca gta  
 cat att ca cac ctt aac act cgc cta ttg tta aag tgt gtc ctt tgt cat  
 M S P I L → GST .....

256 ttc atg tcc cct ata cta ggt tat ttg aaa att aag ggc ctt ttg caa ccc  
 aag tac agg gga tat gat cca ata acc ttt taa ttc ccg gaa cac gtt ggg

919 " GST → R G S R R A S V G S P S T S  
 ctg gtt ccg cgt gga tct cgt cgt gca tct gtt gga tcc cca tca aca agt  
 gac caa ggc gca cct aga gca gca cgt aga caa cct agg ggt agt tgt tca  
 H Y K K aac atg ttt ttc cga cgt gct ctt tcc att ttg cta tat tta tag tta tat



## pDEST3 6823 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
150..200	Trc
1087..963	attR1
1337..1996	CmR
2116..2200	inactivated ccdA
2338..2643	ccdB
2684..2808	attR2
3231..4091	ampR
5295..6254	lacIq

1 ACGTTATCGA CTGCACGGTG CACCAATGCT TCTGGCGTCA GGCAGCCATC GGAAGCTGTG  
 61 GTATGGCTGT GCAGGGCGTA AATCACTGCA TAATTCTGT CGCTCAAGGC GCACCTCCGT  
 121 TCTGGATAAT GTTTTTGCG CCGACATCAT AACGGTTCTG GCAAATATTG TGAAATGAGC  
 181 TGTGACAAT TAATCATCGG CTCGATAATAA GTGTTGAAATT GTGAGCGGAT AACAAATTCA  
 241 CACAGGAAAC AGTATTCTATG TCCCCTATAC TAGGTTATTG GAAAATTAAAG GGCCTTGTGC  
 301 AACCCACTCG ACTTCTTTG GAATATCTTGAAGAAAAATA TGAAGAGCAT TTGTATGAGC  
 361 GCGATGAAGG TGATAAAATGG CGAAACAAAAA AGTTTGAATT GGTTTGGAG TTTCCAATC  
 421 TTCCATTATA TATTGATGGT GATGTTAAAT TAACACAGTC TATGGCCATC ATACGTTATA  
 481 TAGCTGACAA GCACAAACATG TTGGTGTGTT GTCCAAAAGA GCGTGCAGAG ATTCATG  
 541 TTGAAAGGAGC GGTTTTGGAT ATTGATACG GTGTTTCGAG ATTGCTATAT AGTAAAGACT  
 601 TTGAAACTCT CAAAGTTGAT TTCTTAGCA AGCTACCTGA AATGCTGAAA ATGTCGAAG  
 661 ATCGTTTATG TCATAAAAACA TATTTAAATG GTGATCATGT AACCCATCCT GACTTCATGT  
 721 TGTATGACGC TCTTGTGTT GTTTTATACCA TGGACCCAAT GTGCCTGGAT GCGTTCCCAA  
 781 AATTAGTTTG TTTTAAAAAA CGTATTGAAG CTATCCCACA AATTGATAAG TACTTGAAT  
 841 CCAGCAAGTA TATAGCATGG CCTTTGCAGG GCTGGCAAGC CACGTTGGT GGTGGCGACC  
 901 ATCCCTCCAAA ATCGGATCTG GTTCCCGCTG GATCTCGTCG TGATCTGTT GGATCCCCAT  
 961 CAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAATA TGATATAAAAT ATCAATATAT  
 1021 TAAATTAGAT TTTGCATAAA AAACAGACTA CATAATACTG TAAAACACAA CATATCCAGT  
 1081 CACTATGGCG GCCGCTAAGT TGGCAGCATC ACCCGACGCA CTTTGCGCCG AATAAAATACC  
 1141 TGTGACGGAA GATCACTTCG CAGAATAAAAT AAATCCTGGT GTCCCTGTT ATACCGGGAA  
 1201 GCCCTGGGCC AACTTTTGGC GAAAATGAGA CGTTGATCGG CACGTAAGAG GTTCCAACCT  
 1261 TCACCATAAT GAAATAAGAT CACTACCGGG CGTATTTTTT GAGTTATCGA GATTTTCAGG  
 1321 AGCTAAGGAA GCTAAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA  
 1381 ATGGCCTACGT AAAGAACATT TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA  
 1441 GACCGTTCAG CTGGATATTA CGGCCTTTTT AAAGACCGTA AGAAAAATA AGCACAAAGTT  
 1501 TTATCCGGCC TTTATTACACA TTCTTGCCTCG CCTGATGAAT GTCATCCGG AATTCCGTAT  
 1561 GGCAATGAAA GACGGTGAGC TGGTGATATG GGATAGTGT CACCCCTGTT ACACCGTTTT  
 1621 CCATGAGCAA ACTGAAACGT TTTCATCGCT CTGGAGTCAA TACACGACG ATTTCCGGCA  
 1681 GTTCTACAC ATATATTTCG AAGATGTGGC GTGTTACGGT GAAAACCTGG CCTATTTCCC  
 1741 TAAAGGGTTT ATTGAGAATA TGTTTTCTGT CTCAGCCAAT CCCTGGGTGA GTTTCACCAG  
 1801 TTTTGATTTA AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA CCATGGGCAA  
 1861 ATATTATACG CAAGGGGACA AGGTGCTGAT GCCGCTGGCG ATTCAAGGTT ATCATGCCGT  
 1921 CTGTGATGGC TTCCATGTGCG CGACAATGCT TAATGAATTA CAACAGTACT GCGATGAGTG  
 1981 GCAGGGGGGG GCGTAAAGAT CTGGATCCGG CTTACTAAAAA GCCAGATAAC AGTATGGCTA  
 2041 TTTCGCGCT GATTTTTGCG GTATAAGAAT ATATACTGAT ATGTATAACCC GAAGTATGTC  
 2101 AAAAGAGGT GTGCTATGAA GCAGGGTATT ACAGTGACAG TTGACAGCGA CAGCTATCAG  
 2161 TTGCTCAAGG CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC ATGCAGAATG  
 2221 AAGCCCGTCG TCTGCGTGCC GAACGGTGGAA AAGCGAAAAA TCAGGAAGGG ATGGCTGAGG  
 2281 TCGCCCGGTT TATTGAAATG AACGGCTTT TTGCTGACGA GAACAGGGAC TGGTGAATG  
 2341 CAGTTAAGG TTTACACCTA TAAAAGAGAG AGGGTTATC GTCTGTTGT GGATGTACAG  
 2401 AGTGATATTA TTGACACGCC CGGGCGACGG ATGGTGATCC CCGTGGCCAG TGCACGTCTG  
 2461 CTGTCAAGATA AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATGGGGAA TGAAAGCTGG  
 2521 CGCATGATGA CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATGGGGAA AGAAGTGGCT  
 2581 GATCTCAGCC ACCCGGAAAAA TGACATCAA AACGCCATTA ACCTGATGTT CTGGGGATAA  
 2641 TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCGAGTC GACCATAGTG ACTGGATATG-

FIGURE 23B

42/240

2701 TTGTGTTTA CAGTATTATG TAGTCTGTTT TTTATGCAA ATCTAATTAA ATATATTGAT  
 2761 ATTTATATCA TTTTACGTTT CTCGTTCAAGC TTTCCTTGAC AAAGTGGTTG ATGGGAATTG  
 2821 ATCGTGAATG ACTGACGATC TGCCCTCGGC GTTTCGGTGA TGACGGTGA AACCTCTGAC  
 2881 ACATGCAGCT CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGG AGCAGACAAG  
 2941 CCCGTCAGGG CGGGTCAGGG GGTGTTGGCC GGTGTCGGGG CGCAGCCATG ACCCAGTCAC  
 3001 GTAGCGATAG CGGAGTGTAT AATTCTGAA GACGAAAGGG CCTCGTGTATA CGCCTATTTT  
 3061 TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGAA  
 3121 ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAAATATG TATCCGCTCA  
 3181 TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATT  
 3241 AACATTTCGG TGTCGCCCTT ATTCCCCTTT TTGCGGCATT TTGCGCTCTT GTTTTTGCTC  
 3301 ACCCAGAACG GCTGGTGAAGA GTAAAAGATG CTGAAGATCA GTGGGGTGC CGAGTGGGTT  
 3361 ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCG GAAGAACGTT  
 3421 TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG  
 3481 CGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT  
 3541 CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTAA TGCAGTGCTG  
 3601 CCATAACCAC GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA  
 3661 AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG  
 3721 AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA  
 3781 TGGCAACAAC GTTGCACAA CTATTAACGT CGGAACACTA TACTCTAGCT TCCCGCAAC  
 3841 AATTAATAGA CTGGATGGAG CGGGATAAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTC  
 3901 CGGCTGGCTG GTTTATTGCT GATAAACTCG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA  
 3961 TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGA  
 4021 GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA  
 4081 AGCATTGGTA ACTGTCAGAC CAAGTTACT CATATATACT TTAGATTGAT TTAAAACCTC  
 4141 ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTGTA TAATCTCATG ACCAAAATCC  
 4201 CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT  
 4261 CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGC AAAAAAAA CCACCGCTAC  
 4321 CAGCGGTGGT TTGTTTGCCT GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACGGCT  
 4381 TCAGCAGAGC GCAGATAACCA AATACTGTCC TTCTAGTGT GCCGTAGTTA GGCCACCACT  
 4441 TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG  
 4501 CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGACTC AAGACGATAG TTACCGGATA  
 4561 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTG GAGCGAACGA  
 4621 CCTACACCAG ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGGCCACG CTTCCGAAG  
 4681 GGAGAAAGGC GGACAGGTAT CGGTAAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG  
 4741 AGCTTCCAGG GGGAAACGCC TGTTATCTTT ATAGTCTCTGT CGGGTTTCGC CACCTCTGAC  
 4801 TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCACGCA  
 4861 ACGCGGCCCT TTACGGTTC CTGGCTTTT GCTGGCCTTT TGCTCACATG TTCTTCTG  
 4921 CGTTATCCCC TGATTCGTT GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC  
 4981 GCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCTGA  
 5041 TGCGGTATT TCTCCTTACG CATCTGTGCG GTATTTCAACCGCATAAAT TCCGACACCA  
 5101 TCGAATGGTG CAAAACCTT CGCGGTATGG CATGATAGCG CCCGGAAAGAG AGTCAATTCA  
 5161 GGGTGGTGAAT TGTAACCGCA GTAACTGTTT ACAGATGTCG AGAGTATGCC GGTGTCTCTT  
 5221 ATCAGACCGT TTCCCGCGTG GTGAAACCGG CCAGCCACGT TTCTCGGAAA ACGCGGGAAA  
 5281 AAGTGGAAAGC GGCGATGGCG GAGCTGAATT ACATTTCCAA CGCGTGGCA CAACAACCTGG  
 5341 CGGGCAAACAA GTCGTTGCTG ATTGGCGTT CCACCTCCAG TCTGGCCCTG CACGCCCGT  
 5401 CGCAAATTGT CGCGCGATT AAATCTCGC CCGATCAACT GGGTGCCAGC GTGGTGGTGT  
 5461 CGATGGTAGA ACGAAGCGC GTCGAAGCCT GTAAAGCGGC GGTGCACAAAT CTTCTCGC  
 5521 AACCGCGTCAG TGGGCTGATC ATTAACATC CGCTGGATGA CCAGGATGCC ATTGCTGTGG  
 5581 AAGCTGCCTG CACTAATGTT CGCGCGTTT TTCTTGATGT CTCTGACCAAG ACACCCATCA  
 5641 ACAGTATTAT TTTCTCCCAT GAAGACGGTA CGCGACTGGG CGTGGAGCAT CTGGTCGCAT  
 5701 TGGGTACCCA GCAAATCGCG CTGTTAGCGG GCCCATTAAG TTCTGTCTCG GCGCGTCTGC  
 5761 GTCTGGCTGG CTGGCATAAA TATCTCACTC GCAATCAAAT TCAGCCGATA GCGGAACGGG  
 5821 AAGGCGACTG GAGTGCCATG TCCGGTTTTC AACAAACCAT GCAAATGCTG AATGAGGGCA  
 5881 TCGTTCCAC TCGGATGCTG GTTGCCAAAGC ATCAGATGGC GCTGGGCGCA ATGCCGCCA  
 5941 TTACCGAGTC CGGGCTGCCG GTTGGTGCAG ATATCTCGGT AGTGGGATAC GACGATACCG  
 6001 AAGACAGCTC ATGTTATATC CCGCCGTTAA CCACCATCAA ACAGGATT TT CGCCTGCTGG  
 6061 GGCAAACCAAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG CCAGGCGGTG AAGGGCAATC  
 6121 AGCTGTTGCC CGTCTCACTG GTGAAAAGAA AAACCCACCT GGCGCCCAAT ACGCAAACCG-

FIGURE 23C

43/240

6181 CCTCTCCCCG CGCGTTGGCC GATTCAAA TGCAAGCTGGC ACGACAGGTT TCCCGACTGG  
6241 AAAGCGGGCA GTGAGCGCAA CGCAATTAAAT GTGAGTTAGC TCACTCATTA GGCACCCAG  
6301 GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATT  
6361 CACACAGGAA ACAGCTATGA CCATGATTAC GGATTCACTG GCCGTCGTTT TACAACGTCG  
6421 TGACTGGAA AACCCCTGGCG TTACCCAACT TAATGCCCTT GCAGCACATC CCCCTTCGCG  
6481 CAGCTGGCGT AATAGCGAAG AGGCCCCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT  
6541 GAATGGCGAA TGGCGCTTTG CCTGGTTTCC GGCACCCAGAA GCGGTGCCGG AAAGCTGGCT  
6601 GGAGTGCAT CTTCCGTAGG CCGATACTGT CGTCGTCCCC TCAAACGTGGC AGATGCACGG  
6661 TTACGATGCG CCCATCTACA CCAACGTAAC CTATCCCATT ACGGTCAATC CGCCGTTTGT  
6721 TCCCACGGAG AATCCGACGG GTTGTACTC GCTCACATT AATGTTGATG AAAGCTGGCT  
6781 ACAGGAAGGC CAGACGCGAA TTATTTTGA TGGCGTTGGA ATT

FIGURE 23D

44/240

Figure 24A: pDEST<sup>T4</sup>

## His6-thioredoxin fusions in E. coli

919 gca aat att ctg aaa tga gct ggt gac att taa tca tcc ggt ccc **-35 Trc promoter** -10  
 cgt tta taa gac ttt act cga cba ctg tca att agt agg cca ggc ata tca

970 ctg tgg <sup>→ mRN</sup> taat tgt gag egg ata aca att tca cac agg aaa caa acc atg ggc Met **Met**  
 gac acc tta aca ctc gcc tat tgt taa agt gtg tcc ttt gtc egg tac cca

His6

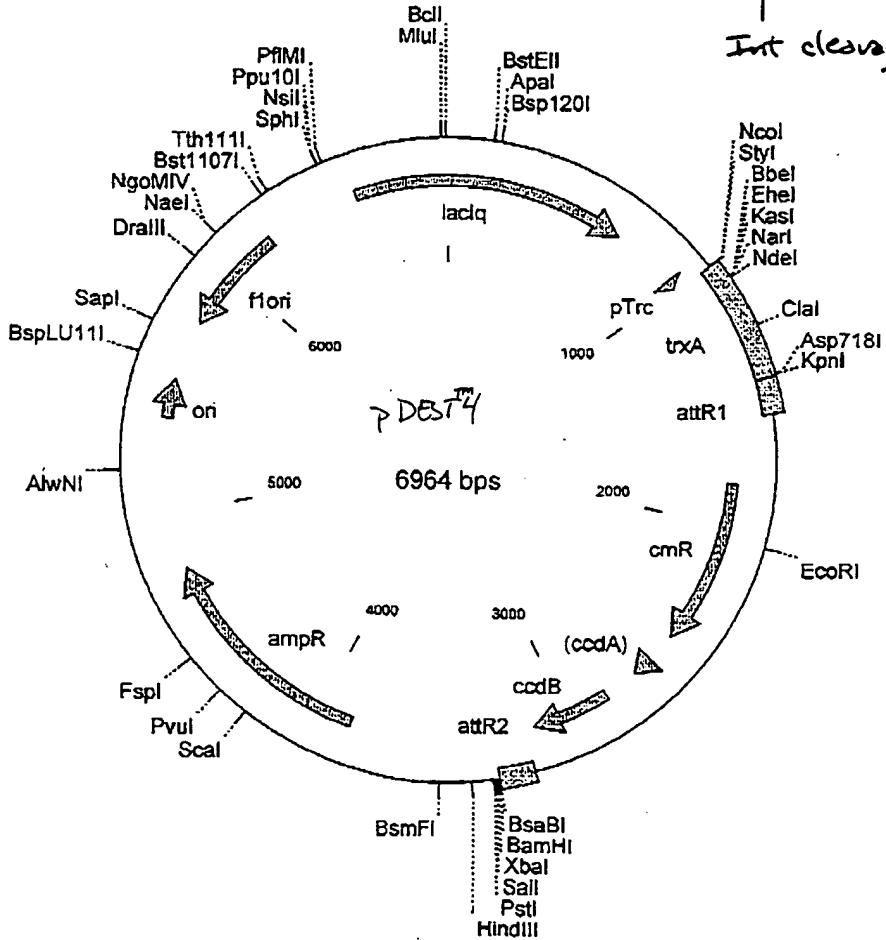
1021 cat cat cat cat His His Tyr Arg Thr Thr His Glu Asp Lys Tyr  
cat cat cat cat cat cac gat tac gat atc cca acg acc gaa aac ctg tde  
gta gta gta gta gta gta gta gtc cta atg cta tag ggt tgc tgg ctt ttg gac ata

TEV protease → Thioredoxin - - (≈ 150 amino acids)  
Phe Gln ↓ Gly Asn His Met Ser Arg Lys Ile Ile His Leu Thr Arg Arg Ser  
tcc cag ggt gcc cat atg agc gat taa att att cac ctg act gac gat agt  
aaa gtc ccc egg gta tac teg cta ttt taa gtg gac tga ctg ctg tca

## attR 1

1429 Gat Gae Gaf Gac Gac Aeg Lys Val Pro Ile Ser Lys Lys Lys  
cta ctg cta ctg ttc cat ggg tag tgg tca aac aro rff lct ega pet gct

Int cleavage



45/260

## pDEST4 6964 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
964..1003	Trc
1577..1453	attR1
1827..2486	CmR
2606..2690	inactivated ccdA
2828..3133	ccdB
3174..3298	attR2
3872..4777	ampR
5378..5538	ori
5778..6215	flori (f1 intergenic region)
6587..704	lacIq

1 CTATCCGCTG GATGACCAGG ATGCCATTGC TGTGGAAAGCT GCCTGCACTA ATGTTCCGGC  
 61 GTTATTTCTT GATGTCTCTG ACCAGACACC CATCAACAGT ATTATTTCTT CCCATGAAGA  
 121 CGGTACGCGA CTGGCGTGG AGCATCTGGT CGCATTGGGT CACCAAGAAA TCGCGCTGTT  
 181 AGCGGGCCCA TTAAGTTCTG TCTCGCGCG TCTCGCTCTG GCTGGCTGGC ATAAATATCT  
 241 CACTCGCAAT CAAATTACAGC CGATAGCGGA ACGGGAAAGGC GACTGGAGTG CCATGTCGG  
 301 TTTTCAACAA ACCATGCAA TGCTGAATGA GGGCATCGTT CCCCACGCGA TGCTGGTTGC  
 361 CAACGATCG ATGGCGCTGG GCGCAATCGC CGCCATTACG GAGTCCGGGC TGCGCGTTGG  
 421 TGCGGATATC TCGGTAGTGG GATAACGACG TACCGAACGAC AGCTCATGTT ATATCCGCC  
 481 GTCAACCACC ATCAAACAGG ATTTTCGCT GCTGGGGCAA ACCAGCGTGG ACCGCTTGCT  
 541 GCAACTCTCT CAGGGCCAGG CGGTGAAGGG CAATCAGCTG TTGCCCCGCT CACTGGTAA  
 601 AAGAAAAAAC ACCCTGGCAC CCAATACGC AACCGCCCTCT CCCCACGCGT TGGCCGATTC  
 661 ATTAATGCAG CTGGCACGAC AGGTTTCCCG ACTGGAAAGC GGGCAGTGAG CGCAACGCAA  
 721 TTAATGTGAG TTAGCGCGA TTGATCTGGT TTGACAGCTT ATCATCGACT GCACGGTGCA  
 781 CCAATGCTTC TGGCGTCAGG CAGCCATCGG AAGCTGTGGT ATGGCTGTGC AGGTCTGAAA  
 841 TCACTGCATA ATTCTGTGTCG CTCAAGGCGC ACTCCGTTT TGGATAATGT TTTTIGGCC  
 901 GACATCATAA CGGTTCTGGC AAATATTCTG AAATGAGCTG TTGACAATTAA ATCATCCGGT  
 961 CCGTATAATC TGTGGAATTG TGAGCGGATA ACAATTTCAC ACAGGAAACA GACCATGGGT  
 1021 CATCATCATC ATCATCACCA TTACGATATC CCAACGACCC AAAACCTGTA TTTTCAGGGC  
 1081 GCCCATATGA GCGATAAAAT TATTCACCTG ACTGACGACA GTTTTGACAC GGATGTACTC  
 1141 AAAGCGGACG GGGCGATCCT CGTCGATTTC TGGGCAGAGT GGTGCGGTCC GTGAAAATG  
 1201 ATCGCCCCGA TTCTGGATGA AATCGCTGAC GAATATCAGG GCAAACGTAC CGTTGAAAAA  
 1261 CTGAACATCG ATCAAACCC TGGCACTGCG CCGAAATATG GCATCCGTG TATCCGACT  
 1321 CTGCTGCTGT TCAAAAACGG TGAAGTGGCG GCAACCAAAG TGGGTGCACT GTCTAAAGGT  
 1381 CAGTTGAAAG AGTTCCCTCGA CGCTAACCTG GCCGGTTCTG GTTCTGGTGA TGACGATGAC  
 1441 AAGGTACCCA TCACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT  
 1501 ATCAATATAT TAAATTAGAT TTTGATCTAA AAACAGACTA CATAATACTG TAAAACACAA  
 1561 CATATCCAGT CACTATGGCG GCCGCTAAGT TGGCAGCATC ACCCGACCCA CTTTGCACCG  
 1621 AATAAAATACC TGTGACGGAA GATCACTTCG CAGAATTAAT AAATCTGTG GTCCCTGTG  
 1681 ATACCGGGAA GCCCTGGGGC AACTTTGGC GAAAATGAGA CGTTGATCGG CACGTAAGAG  
 1741 GTTCCAACCT TCACCATATAA GAAATAAGAT CACTACCGG CGTATTTTTT GAGTTATCGA  
 1801 GATTTTCAGG AGCTAAGGAA GCTAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG  
 1861 ATATATCCCA ATGGCATCTG AAAGAACATT TTGAGGGATT TCAGTCAGTT GCTCAATGTA  
 1921 CCTATAACCA GACCGTTCACTG CTGGATATTA CGGCCTTTTTT AAAGACCGTA AAGAAAAATA  
 1981 AGCACAAGTT TTATCCGGGTTT TTTATTACCA TTCTTGCCCG CCTGATGAAT GCTCATCCGG  
 2041 AATTCGTAT GGCAATGAAA GACGGTGAGC TTGGATGATATG GGATAGTGT CACCCCTGTT  
 2101 ACACCGTTTT CCATGAGCAA ACTGAAACAGT TTTCATCGCT CTGGAGTGA TACCAACGAC  
 2161 ATTTCCGGCA GTTCTACAC ATATATTCTG AAGATGTGGC GTGTTACGGT GAAAACCTGG  
 2221 CCTATTTCCC TAAAGGGTTT ATTGAGAATA TTGTTTTCTG CTCAGCCAAAT CCCTGGGTGA  
 2281 GTTTCACCGAG TTTTGTATTA AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA  
 2341 CCATGGGCAA ATATTATACG CAAGGGCACA AGGTGCTGAT GCGCGTGGG ATTCAAGGTT  
 2401 ATCATGCCGT CTGTGATGGC TTCCATGTGCG GCAGAAATGCT TAATGAATTAA CAACAGTACT  
 2461 GCGATGAGTG GCAGGGCGGG GCGTAAACGC GTGGATCCGG CTTACTAAA GCCAGATAAC  
 2521 AGTATGCGTA TTTGCGCGCT GATTTTGCGC GTATAAGAAT ATATACTGAT ATGTATACCC-

FIGURE 24B

46/240

2581 GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT ACAGTGACAS TTGACAGCGA  
 2641 CAGCTATCG TTGCTCAAGG CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC  
 2701 ATGCAGAATG AAGCCCGTCG TCTGCGTGC GAACGCTGGA AAGCGGAAAA TCAGGAAGGG  
 2761 ATGGCTGAGG TCGCCCCGGT TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC  
 2821 TGTTGAAATG CAGTTAAGG TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT  
 2881 GGATGTACAG AGTGTATTTA TTGACACGCC CGGGCGACGG ATGGTGTACCC CCCTGGCCAG  
 2941 TGCACGCTG CTGTCAGATA AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATCGGGG  
 3001 TGAAAGCTGG CGCATGATGA CCACCGATAT GGGCAGTGTG CCGGTCTCCG TTATCGGGG  
 3061 AGAAGTGGCT GATCTCAGCC ACCGCGAAAA TGACATCAA AACGCCATTA ACCTGATGTT  
 3121 CTGGGAAATA TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCAAGTC GACCATACTG  
 3181 ACTGGATATG TTGTTGTTTA CAGTATTATG TAGTCTGTTT TTTATGCAAATCTAATTTA  
 3241 ATATATTGAT ATTTATATCA TTTTACGTTT CTGCTTCAGC TTTCTTGAC AAAGTGGTGA  
 3301 TGGGGATCCT CTAGACTCGA CCTGCACTAA TCGTACAGGG TAGTACAAAT AAAAAGGCA  
 3361 CGTCAGATGA CGTGCCTTTT TTCTTGTGAG CAGTAAGCTT GGCTGTTTG GCGGATGAGA  
 3421 GAAGATTTTC AGCCTGATAC AGATTAATAC AGAACCGAGA AGCGGTCTGA TAAAACAGAA  
 3481 TTTGCTGGC GGCACTAGCG CGGTGGTCCC ACCTGACCC ATGCCGAAC CAGAAGTGA  
 3541 ACGCCGTAGC GCCGATGGTA GTGTGGGTC TCCCCATGCC AGACTAGGCA ACTGCCAGGC  
 3601 ATCAAATAAA ACAGAAAGGCT CAGTCGAAAG ACTGGGCCTT CGTTTTATC TGTTGTTTGT  
 3661 CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCGGGAGC GGATTTGAAC GTTGCAGAAGC  
 3721 AACGGCCCGG AGGGTGGCGG GCAGGACGCC CGCCATAAAC TGCCAGGAT CAAATTAAAGC  
 3781 AGAAGGCCAT CCTGACGGAT GGCCTTTTG CGTTTCTACA AACTCTTTT GTTTATTTT  
 3841 CAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCATA  
 3901 ATATTGAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTT  
 3961 TGGGGATTT TGCCCTTCCTG TTTTGCTCA CCCAGAAACG CTGGTAAAG TAAAAGATGC  
 4021 TGAAGATCAG TTGGGTGAC GAGTGGGTTA CATCGAACTG GATCTAACAA CGGTTAAGAT  
 4081 CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAAATGATG AGCACTTTA AAGTTCTGCT  
 4141 ATGTGGCGCG GTATTATCCC GTGTTGACGC CGGGCAAGAG CAACTCGTC GCGCATAACA  
 4201 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG  
 4261 CATGACAGTA AGAGAATTAT GCAGTCTGC CATAACCAGT AGTGATAACA CTGCGCCAA  
 4321 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACCG GCTTTTTGCAAC  
 4381 GGATCATGTA ACTCGCCTTG ATCGTTGGG ACCGGAGCTG AATGAAGCCA TACCAACAGA  
 4441 CGAGCGTGC ACCACGATGC CTACAGCAAT GGCAACAAACG TTGCGCAAAC TATTAACTGG  
 4501 CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAAGT  
 4561 TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTATGCTG ATAAATCTGG  
 4621 AGCCGGTGG AGTGGGTCTC GCGGTATCAT TGAGCTACTG GGGCCAGATG GTAAGCCCTC  
 4681 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA  
 4741 GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTACTC  
 4801 ATATATACTT TAGATTGATT TAAAACCTCA TTTTAATTT AAAAGGATCT AGGTGAAGAT  
 4861 CCTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC  
 4921 AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGTACCT TTTTTCTGC GCGTAATCTG  
 4981 CTGCTTGCAA ACACAAAAAC CACCGCTACC AGCGGTGGTT TGTGCGCCG ATCAAGAGCT  
 5041 ACCAACTCTT TTTCGAAAGG TAACTGCTT CAGCAGAGCG CAGATACCAA ATACTGTCCT  
 5101 TCTAGTGTAG CGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT  
 5161 CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCT GTCTTACCGG  
 5221 GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTT  
 5281 GTGCACACAG CCCAGCTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA  
 5341 GCTATGAGAA AGCGCCACCG TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG  
 5401 CAGGGTCGGA ACAGGAGAGC GCACGAGGG GCTCCAGGG GGAAACGCT GGTTACTTTA  
 5461 TAGTCTGTC GGGTTTCGCC ACCTCTGACT TGAGCTGCA TTTTGTTGAT GTCGTCAGG  
 5521 GGGCGGGAGC CTATGAAAAA ACACCGCAA CGCCGCCCTT TTACGGGTTCC TGGCCTTTG  
 5581 CTGGCCTTT GCTCACATGT TCTTCTCTG GTTATCCCT GATTCTGTTG ATAACCGTAT  
 5641 TACCGCCTT GAGTGTGAGCTG ATACCGCTCG CCGCAGCGGA ACAGACCGAGC GCAGCGAGTC  
 5701 AGTGTGAGCAG GAAGCGGAAG AGCGCCTGAT GCGGTATTT CTCCCTACGC ATCTGTGCGG  
 5761 TATTTACAC CGCATAATTG TGTTAAAATT CGCGTTAAAT TTTTGTAAA TCAGCTCATT  
 5821 TTTTAACCAA TAGGCGAAA TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT  
 5881 AGGGTTGAGT GTTGTGTCAG TTTGGACCAA GAGTCCACTA TAAAGAACG TGGACTCCAA  
 5941 CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCTA  
 6001 ATCAAGTTTT TTGGGGTCAA GGTGCGTAA AGCACTAAAT CGGAACCCCTA AAGGGAGCCC-

FIGURE 24C

47/240

6061 CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC  
6121 GAAAGGAGCG GGCCCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TA-ACCACAC  
6181 ACCCGCCGCG CTTAATGCC CGCTACAGGG CGCGTCCATT CGCCATTCAAG GCTGCTATGG  
6241 TGCACCTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGTATAAC ACTCCGCTAT  
6301 CGCTACGTGA CTGGGTCACTG GCTGCCGCCCC GACACCCGCC AACACCCGCT GACGCCCT  
6361 GACGGGCTTG TCTGCTCCC GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT  
6421 GCATGTGTCA GAGGTTTCA CCGTCATCAC CGAACGCGC GAGGCAGCAG ATCAATTGCG  
6481 GCGCGAAGGC GAAGCGGCAT GCATTTACGT TGACACCATC GAATGGTGCA AAACCTTCG  
6541 CGGTATGGCA TGATAGCGCC CGGAAGAGAG TCAATTCAAG GTGGTGAATG TGAAACCAAGT  
6601 AACGTTATAC GATGTCGCAAG AGTATGCCGG TGTCTCTTAT CAGACCGTTT CCCGCGTGGT  
6661 GAACCAGGCC AGCCACGTTT CTGCGAAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA  
6721 GCTGAATTAC ATTCCCAACC GCGTGGCAACA ACAACTGGCG GGCAAACAGT CGTTGCTGAT  
6781 TGGCGTTGCC ACCTCCAGTC TGGCCCTGCA CGCGCCGTCG CAAATTGTCG CGGCATTAA  
6841 ATCTCGCGCC GATCAACTGG GTGCCAGCGT GGTGGTGTGCG ATGGTAGAAC GAAGCGGCGT  
6901 CGAAGCCTGT AAAGCGGCGG TGCACAATCT TCTCGCGCAA CGCGTCAGTN GGGCTGATCA  
6961 TTAA

FIGURE 24D

48/240

Figure 25A

pDEST5

pSPORT '+' (for sequencing, probes,  
phagemid)

1. agg cac ccc agg ~~cct tac act tta tgc ttc cgg ctc gta tgt tpt~~ lac promoter -10 lac RNA  
 tcc gtg ggg tcc ~~gaa atg tga aat acg aag gcc gag cat aca gca cac ctt~~

"reverse" sequencing primers

52 ttg tga gcg gat aac aat ttc aca cag gaa aca get <sup>→ α- peptide</sup> aac act cgc cta ttg tta aag tgt gtc ctt tgt cga tac tgg tac taa tgc

103 cca agc ~~tct aat acg act cac tat agg gaa agc tgg tac gcc tgc~~ T7 promoter T7 RNA Pst Kpn  
 ggt tcc agt ~~tta tgc tga gtg ata tcc gtt tcc acc atg cgg acg tcc atg~~

154 EcoRI Sma Sst I Int attR1  
 cgg tcc ~~ggg att ccc ggg tcg acg atc aca agt tgg xac xaa gaa gct gaa~~  
 gcc agg cct taa ~~ggg ccc agc tgc tag tgt tca aac atg ttt tcc cga ott~~

Gene

Int attR2

1990 ~~tct acg ttt ctc gtt cag ctc tct tgt aca aag tgg tga tca lcta gtc ggc~~ Spe  
 aaa cgc xaa gaa caa gtc gaa aga Aca ~~tgt ttc acc act agt gat dag ccc~~

2041 Not Xba Bam Hind3 Mlu Sph  
~~ggc cgc tct aga gga tcc aag ctt acg tac gca tgc atg cga cgt cat agc~~  
 ccc ggg aga tct cct agg ttc gaa tgc atg cgc aca tac gct gca gta tcc

2092 tct ~~tct ata gtg tca cct aaa ttc aat tca ctg gcc gtc gtt tta caa cgt~~ SP6 promoter  
 aga ~~aga tat cac agt gga ttt aag tta agt gac egg cag caa aat gtt gca~~  
 ← SP6 RNA

"forward sequencing . . ."

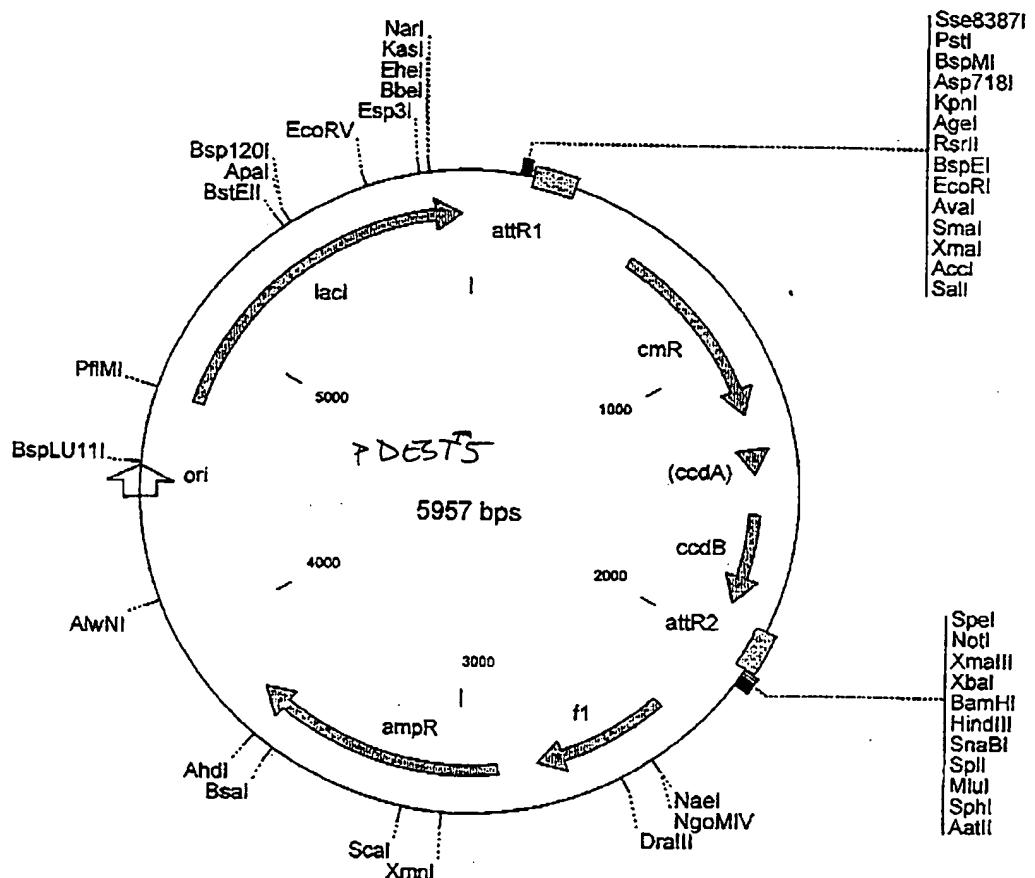
2143 cgt gac tgg gaa aac cct ggc gtt acc caa ctt aat cgc ctt gca gca cat  
 gca ctg acc ctt ttg gga ccg ~~caa tgg gtt gaa tta gcg gaa cgt cgt gta~~  
 .. primers

49/240

Figure 25B

 $\lambda$  DESTS

(cont'd)



50/240

## pDEST5 5957 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
305..181	attR1
555..1214	CmR
1334..1418	inactivated ccdA
1556..1861	ccdB
1902..2026	attR2
2278..2733	f1 (f1 intergenic region)
2865..3722	ampR
5378..5538	ori
4756..5922	lacI

1 AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG  
 61 GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTC TAATACGACT  
 121 CACTATAGGG AAAGCTGGTA CGCCTGCAGG TACCCGGTCCG GAATTCCCGG GTGCACGATC  
 181 ACAAGTTGT ACAAAAAAGC TGAACCAGAA ACCTAAAATG ATATAAAATAT CAATATATTA  
 241 AATTAGATTT TGCACTAAAA ACAGACTACA TAATACTGTA AAACACAAACA TATCCAGTCA  
 301 CTATGGCGGC CGCTAAGTTG GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG  
 361 TGACGGAAGA TCACTTCGCA GAATAAAATAA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC  
 421 CCTGGGCCAA CTTTGGCGA AAATGAGACG TTGATCGGCA CGTAAGAGGT TCCAACTTTC  
 481 ACCATAATGA AATAAGATCA CTACGGGGCG TATTTTTGTA GTTATCGAGA TTTTCAGGAG  
 541 CTAAGGAAGC TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT  
 601 GGCATCGTAA AGAACATTTC GAGGCATTTC AGTCAGTTGC TCAATGTACC TATAACCAGA  
 661 CCGTTCAGCT GGATATTACG GCCTTTTAA AGACCGTAAAG GAAAAATAAG CACAAGTTT  
 721 ATCCGGCCTT TATTCACATT CTTGCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG  
 781 CAATGAAAGA CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTTGTAC ACCGTTTCC  
 841 ATGAGCAAAC TGAAACGTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT  
 901 TTCTACACAT ATATTGCAAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA  
 961 AAGGGTTAT TGAGAATATG TTTTTCGCTC CAGCCAATCC CTGGGTGAGT TTCACCAAGTT  
 1021 TTGATTTAAA CGTGGCCAAT ATGGACAATCT TCTTCGCCCC CGTTTTCAAC ATGGGCAAAT  
 1081 ATTATACGCA AGGCAGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT  
 1141 GTGATGGCTT CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC  
 1201 AGGGGGGGC GTAAAAGCGT GGATCCGGT TACTAAAAGC CAGATAACAG TATGCGTATT  
 1261 TGCGCCTGA TTTTGCCTG ATAAGAATAT ATACTGATAT GTATACCCGA AGTATGTCAA  
 1321 AAAGAGGTGT GCTATGAAGC AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT  
 1381 GCTCAAGGCA TATATGATGT CAATATCTCC GGTCTGGTAA GCACAAACCAT GCAGAATGAA  
 1441 GCCCGTCGTC TGCGTGGCGA ACGCTGGAAA CGGAAATAC AGGAAGGGAT GGCTGAGGTC  
 1501 GCCCGTTTA TTGAAATGAA CGGCTTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA  
 1561 GTTAAGGTT TACACCTATA AAAGAGAGAG CGCTTATCGT CTGTTTGTGG ATGTACAGAG  
 1621 TGATATTATT GACACGCCCG GGCGACGGAT GGTGATCCCC CTGGCCAGTG CACGCTCTGCT  
 1681 GTCAGATAAA GTCTCCCGTG AACTTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG  
 1741 CATGATGACC ACCGATATGG CCAGTGTGCC GGTCTCCGTT ATCGGGGAAAG AAGTGGCTGA  
 1801 TCTCAGGCCAC CGCGAAAATG ACATCAAAAA CGCCATTAAAC CTGATGTTCT GGGGAATATA  
 1861 AATGTCAGGC TCCCTTATAC ACAGCCAGTC TGCAAGTCGA CCATAGTGCAC TGGATATGTT  
 1921 GTGTTTACA GTATTATGTA GTCTGTTTT TATGCAAAT CTAAATTAAAT ATATTGATAT  
 1981 TTATATCATT TTACGTTCT CGITCAGCTT TCTTGTACAA AGTGGTGCATC ACTAGTCGGC  
 2041 GGCCGCTCTA GAGGATCCAA GCTTACGTAC CGGTGCTATGC GACGTCATAG CTCTTCTATA  
 2101 GTGTCACCTA AATTCAATTG ACTGGCCGTC GTTTTACAAAC GTCTGACTG GGAAAACCT  
 2161 GGCCTTACCC AACTTAATCG CCTTGAGCA CATCCCCCTT TCGCCAGCTG GCGTAATAGC  
 2221 GAAGAGGCCCG GCACCGATCG CCCTTCCCAA CAGTTGCCA GCCTGAATGG CGAATGGACG  
 2281 CGCCCTGTAG CGGCGCATTA AGCGCCGGCG GTGTGGTGGT TACGCGCAGC GTGACCGCTA  
 2341 CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCTTCCCTT CTCGCCACGT  
 2401 TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTAGGGTTC CGATTAGTGC  
 2461 CTTTACGGCA CCTCGACCCCA AAAAAGCTTG ATTAGGGTGA TGTTTCACGT AGTGGGCCAT  
 2521 CGCCCTGATA GACGGTTTT CGCCCTTGA CGTGGAGTC CACGTTCTT AATAGTGGAC  
 2581 TCTTGTCCA AACTGGAACA AACTCAACC CTATCTCGGT CTATTCTTT GATTATAAG

FIGURE 25C

51/240

2641 GGATTTGCC GATTCGGCC TATTGGTAA AAAATGAGCT GATTTAACAA AAATTTAAC  
 2701 CGAATTTAA CAAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTICCG GGAAATGTGC  
 2761 GCGGAACCCC TATTTGTTA TTTTTCTAAA TACATTCAA TATGTATCCG CTCATGAGAC  
 2821 AATAACCCCTG ATAAATGTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT  
 2881 TCCGTGTCGC CCTTATTCCC TTTTTGCGG CATTTCGCT TCCTGTTTT GCTCACCCAG  
 2941 AACCGCTGGT GAAAGTAAA GATGCTGAAG ATCAGTTGG TGACAGAGTG GGTTACATCG  
 3001 AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTCG CCCCGAAGAA CGTTTCCAA  
 3061 TGATGAGCAC TTTTAAAGT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGC  
 3121 AAGAGCAACT CGGTGCGCCG ATACACTATT CTACAATGA CTTGGTTGAG TACTCACCAG  
 3181 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA  
 3241 CCATGAGTGA TAACACTGGC GCCAACCTAC TTCTGACAAC GATCGGAGGA CCGAAGGAGC  
 3301 TAACCGCTTT TTTGACAAAC ATGGGGGATC ATGTAACCTCG CCTTGATCGT TGGGAACCGG  
 3361 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCCAC GATGCCTGTA GCAATGGCAA  
 3421 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA  
 3481 TAGACTGGAT GGAGGGGGAT AAAGTTGCGAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG  
 3541 GCTGGTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCAITGCAG  
 3601 CACTGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG  
 3661 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCTCTACTG ATTAAGCATT  
 3721 GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAA CTTCATTTT  
 3781 AATTAAAAG GATCTAGGT AAGATCCTT TTGATAATCT CATGACCAAA ATCCCTTAAC  
 3841 GTGAGTTTTC GTTCACTGA GCGTCAGACC CCGTAGAAA GATCAAAGGA TCTTCTTGAG  
 3901 ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCACCG CTACCAGCGG  
 3961 TGGTTTGTGTT GCGGATCAA GAGCTACCAA CTCTTTTCC GAAGGTAAC GGCTTCAGCA  
 4021 GAGCGCAGAT ACCAAATACT GTCCTCTAG TGTAGCCGTG GTTAGGCCAC CACTTCAAGA  
 4081 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAAGTG GCTGCTGCCA  
 4141 GTGGCGATAA GTCGTGCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGC  
 4201 AGCGGTCGGG CTGAACGGGG GGTTCTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA  
 4261 CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA  
 4321 AGGGGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACACGG AGAGCGACG AGGGAGCTTC  
 4381 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGTT TCGCCACCTC TGACTTGAGC  
 4441 GTCGATTTTT GTGATGCTCG TCAGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCC  
 4501 CCTTTTTACG GTTCCCTGCC TTTTGCTGGC CTTTTGCTCA CATGTTCTTT CCTGCGTTAT  
 4561 CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGCT AGCTGATACC GCTGCCGCC  
 4621 GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGGAGGAAGC GGAAGAGCGC CCAATACCCA  
 4681 AACCGCCTCT CCCC CGCGCGT TGGCGATTG ATTAATGCAG AGCTTGCAAT TCGCGCGCA  
 4741 AGGGGAAGCG GCATTTACGT TGACACCACAT GAATGGCGA AAACCTTTCG CGGTATGGCA  
 4801 TGATAGCGCC CGGAAGAGAG TCAATTTCAGG GTGGTGAATG TGAAACCACT AACGTTATAC  
 4861 GATGTCGCAG AGTATGCGG TGTCTCTTAT CAGACCGTTT CCGCGTGGT GAAACAGGCC  
 4921 AGCCACGTTT CTGCGAAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA GCTGAATTAC  
 4981 ATTCCCAACCG CGGTGGCACA ACAAATGGCG GGCAAAACAGT CGTTGCTGAT TGGCGTTGCC  
 5041 ACCTCCAGTC TGGCCCTGCA CGCGCCGTG CAAATTGTCG CGGCGATTAA ATCTCGCGCC  
 5101 GATCAACTGG GTGCCAGCGT GGTGGTGTG ATGGTAGAAC GAAGCGCGT CGAACCTGT  
 5161 AAACGGCGG TGCACAAATCT TCTCGCGAA CGGGTCAGTG GGCTGATCAT TAACTATCCG  
 5221 CTGGATGACC AGGATGCCAT TGCTGTGGAA GCTGCCTGCA CTAATGTTCC GGGTTATTT  
 5281 CTTGATGTC CTGACCGAC ACCCATCAAC AGTATTATT TCTCCCATGA AGACGGTAGC  
 5341 CGACTGGCG TGGAGCATCT GGTGCATTG GGTCAACAGC AAATCGCGT GTTACCGGGC  
 5401 CCATTAAGTT CTGTCCTGGC GCGCTCGCGT CTGGCTGGCT GGCATAAAATA TCTCACTCGC  
 5461 AATCAAATTC AGCGGATAGC GGAAACGGGAA GGCGACTGGA GTGCCATGTC CGGTTTCAA  
 5521 CAAACCATGC AAATGCTGAA TGAGGGCATC GTTCCCACTG CGATGCTGGT TGCCAACGAT  
 5581 CAGATGGCGC TGGCGCAAT GCGGCCATT ACCGAGTCCC GGCTGCCGT TGGTGCAGGAT  
 5641 ATCTCGGTAG TGGGATAACGA CGATACCGAA GACAGCTCAT GTTATATCCC GCGGTCAACC  
 5701 ACCATCAAAC AGGATTTTCG CCTGCTGGGG CAAACCAAGCG TGGACCGCTT GCTGCAACTC  
 5761 TCTCAGGGCC AGGCGGTGAA GGGCAATCAG CTGTTGCCG TCTCACTGGT GAAAAGAAAA  
 5821 ACCACCCCTGG CGCCCAATAC GCAAACCGCC TCTCCCGCG CGTTGCCGA TTCATTAATG  
 5881 CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCAGT GAGCGCAACG CAATTAATGT  
 5941 GAGTTAGCTC ACTCATT

FIGURE 25D

52/240

Figure 26A pDEST6

pSPORT " " (opposite strand)

## "forward" sequencing primers

1 taa/cgc cag ggt ttt ccc agt cac gac gtt gta aaa cga cgg cca gtg aat  
att gcg gtc cca aaa ggg tca gtg ctg caa cat ttt gct gcc ggt cac tta

52 SP6 promoter Sph Mlu  
tga att tag gty aca cta tag aag agc tat gac gtc gca tgt aeg cgt aeg  
act taa atc cac tgt atc ttc tgc ata ctg cag gtt aeg tgc gca tgc

103 Hind 3 Bam Xba Not Spe att R1 Int  
tata act tag atc ctc tag agc lggc cgc cga tta gtg atc aca agt tgg paf  
att cga acc tag gag atc tcc ccc gtc gct gat dac tag tgt tca aac atc

154 aaa daa get gaa cga gaa acg tax aat gar ata aat acc aat ata red aat  
ttt tet cga ctt get ctt tge att tta cta tat tca tag tta tat aat tca

↓ Gene

1939 Int att R2  
tag tta tat tat tt acg att ctc gtt tag crt pct tgc aca aag tgg tga  
cta dat ata gta aaa gtc aac gag eaa gtc gaa aga aca gtc ttc acc aat

1990 Sal Sma EcoRI Kpn Pst  
tcg tcc acc cgg daa ttc cgg acc ggt acg tgc agg cgt acc agc ttt ccc  
agc agg gcc ctt aag gcc tgg dca tgg acg tcc gca tgg tcc aaa ggg  
T7 RNA

2041 tat agt gag tcg tat tag age ttg gcg taa tca tgg tca tag ctg ttt cct  
ata tca ctc agc ata atc tgc aac cgc att agt acc agt atc gac aaa gga  
T7 promoter α-peptide

"reverse .."

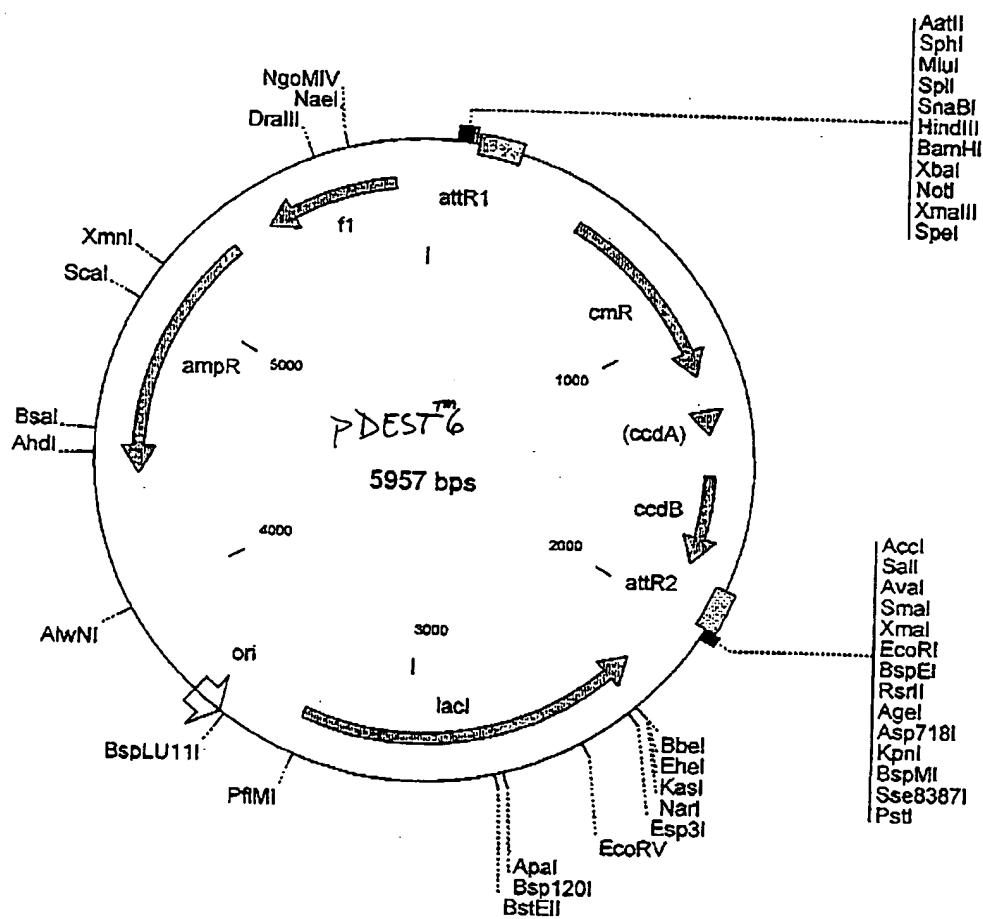
2092 gtg tga aat tgt tat ccg ctc aca att cca cac <sup>-10</sup> lac promoter  
cac act tta aca ata ggc gag tgt taa ggt gtg tgc ttt gca gct gga agc  
... sequencing primers lac RNA

2143 <sup>-35</sup> ata aag tgt aaa gcc tgg ggt gcc taa tga gtg agc taa ctc aca tta att  
tat ttc aca ttc cgg acc cca cgg att act cac tcc att gag tgt aat taa

53/240

**Figure 26B**

PDEST6 (cont'd)



54/240

## pDEST6 5957 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
266..142	attR1
516..1175	CmR
1295..1379	inactivated ccdA
1517..1822	ccdB
1863..1987	attR2
2203..3369	lacI
4403..5260	ampR
5392..5847	f1 (f1 intergenic region)

1 TAACGCCAGG GTTTTCCAG TCACGACGTT GTAAAACGAC GGCCAGTGAA TTGAATTTAG  
 61 GTGACACTAT AGAACAGCTA TGACGTCGCA TGACACCGTA CGTAAGCTTG GATCCTCTAG  
 121 AGCGGCCGCC GACTAGTGTG CACAAGTTG TACAAAAAAAG CTGAACGAGA AACGTAAAAT  
 181 GATATAAATA TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT  
 241 AAAACACAAAC ATATCCAGTC ACTATGGCGG CCGCTAAGTT GGCAGCATCA CCCGACGCAC  
 301 TTTGCGCCGA ATAAAATACCT GTGACGGAAG ATCACTTCGC AGAATAAAATA AATCCTGGTG  
 361 TCCCTGTTGA TACCGGGAAAG CCCTGGGCCA ACTTTTGGCG AAAATGAGAC GTTGATCGGC  
 421 ACGTAAAGAGG TTCCAACCTT CACCATAATG AAATAAGATC ACTACCGGGC GTATTTTTTG  
 481 AGTTATCGAG ATTTTCAGGA GCTAAGGAAG CTAAAATGGA GAAAAAAATC ACTGGATATA  
 541 CCACCGTTGA TATATCCCAA TGGCATCGTA AAGAACATT TGAGGCATTT CAGTCAGTTG  
 601 CTCATGTAC CTATAACCAG ACCGTTCAAGC TGGATATTAC GGCCCTTTTA AAGACCGTAA  
 661 AGAAAAATAA GCACAAGTTT TATCCGGCT TTATTACAT TCTTGGCCGC CTGATGAATG  
 721 CTCATCCGGA ATTCCGTATG GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC  
 781 ACCCTTGTGA CACCGTTTTC CATGAGCAAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT  
 841 ACCACGACGA TTTCCGGCAG TTTCTACACA TATATTGCA AGATGTGGCG TGTTACGGTG  
 901 AAAACCTGGC CTATTTCCCT AAAGGTTTA TTGAGAATAT GTTTTCTGTC TCAGCCAATC  
 961 CCTGGGTGAG TTTCACCAAGT TTTGATTAA ACGTGGCCAA TATGGACAAC TTCTTCGCC  
 1021 CCGTTTTCAC CATGGGCAAA TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA  
 1081 TTCAGGTTCA TCATGCCCTC TGTGATGGCT TCCATGTCGG CAGAATGCTT AATGAATTAC  
 1141 AACAGTACTG CGATGATGG CAGGGCGGGG CGTAAACGCG TGGATCCGGC TTACTAAAAG  
 1201 CCAGATAACA GATACGCTAT TTGCGCGCTG ATTTTGCGG TATAAGAATA TATACTGATA  
 1261 TGTATACCCG AAGTATGTC AAAAGAGGTG TGCTATGAAAG CAGCGTATTAA CAGTGACAGT  
 1321 TGACAGCGAC AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA  
 1381 AGCACAACCA TGCGAAATGA AGCCCGTCGT CTGCGTGGCG AACGCTGGAA AGCGGAAAAT  
 1441 CAGGAAGGGA TGGCTGAGGT CGCCCGGTTT ATTGAAATGA ACGGCTCTT TGCTGACGAG  
 1501 AACAGGGACT GGTGAAATGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG  
 1561 TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCC GGGCGACGGA TGGTGATCCC  
 1621 CCTGGCCAGT GCACGTCTGC TGTCAAGATA AGTCTCCCGT GAAACTTACCG CGGTGGTGCA  
 1681 TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT  
 1741 TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAA ACGCCATTAA  
 1801 CCTGATGTTG TGGGGATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTG  
 1861 ACCATAGTGA CTGGATATGT TGTGTTTAC AGTATTATGT AGTCTGTTT TTATGCAAA  
 1921 TCTAATTTAA TATATTGATA TTTATATCAT TTTACGTTTC TCGTTCAGCT TTCTTGTACA  
 1981 AAGTGGTGTAT CGTCGACCCG GGAATTCCGG ACCGGTACCT GCAGGGTAC CAGCTTCCC  
 2041 TATACTGAGT CGTATTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTC TGTGTAAT  
 2101 TGTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAAGTG TAAAGCCTGG  
 2161 GGTGCCTAAT GAGTGGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTCCAG  
 2221 TCGGGAAACC TGCGTGCAC GCTGCATTAA TGAATCGGCC AACGGCGGG GAGAGGCGGT  
 2281 TTGCGTATTG GGCGCCAGGG TGGTTTTCT TTTCACCAAGT GAGACGGGCA ACAGCTGATT  
 2341 GCCCTTCACC GCCTGGCCCT GAGAGAGTTG CAGCAAGCGG TCCACGCTGG TTTGCCCG  
 2401 CAGGGAAAAA TCCTGTTGA TGGTGGTTGA CGGGGGATA TAACATGAGC TGTCTTCGGT  
 2461 ATCGTCGTAT CCCACTACCG AGATATCCGC ACCAACGCGC AGCCCGGACT CGGTAATGGC  
 2521 GCGCATTGCG CCCAGCGCCA TCTGATGTT GGCAACCAGC ATCGCAGTGG GAACGATGCC  
 2581 CTCATTCAAGC ATTTGCATGG TTTGTTGAAA ACCGGACATG GCACTCCAGT CGCCTTCCCG  
 2641 TTCCGCTATC GGCTGAATTG GATTGCGAGT GAGATATTAA TGCCAGGCCAG CCAGACGCAG-

FIGURE 26C

2701 ACCGGCCGAG ACAGAACTTA ATGGGCCGCG TAACAGCCG ATTTGCTGGT GACCCAATGC  
 2761 GACCAAGATGC TCCACGCCA GTCGCGTACC GTCTTCATGG GAGAAAATAA TACTGTTGAT  
 2821 GGGTGTCTGG TCAGAGACAT CAAGAAATAA CGCCGGAAAC TTAGTGCAGG CAGCTTCCAC  
 2881 AGCAATGGCA TCCTGGTCAT CCAGGGATA GTTAATGATC AGCCCACGTGAC CCCGTTGCGC  
 2941 GAGAAAGATTG TGCAACGCCG CTTTACAGGC TTCGACGCCG CTTCGTTCTA CCATCGACAC  
 3001 CACCAACGCTG GCACCCAGTT GATCGCGCG AGATTTAAC GCGCGACAA TTTGCGACGG  
 3061 CGCGTGCAGG GCCAGACTGG AGGTGGCAAC GCGAACATCAG AACGACTGTT TGCCCGCCAG  
 3121 TTGGTGTGCC ACAGCGTTGG GAATGTAATT CAGCTCCGCC ATCGCCGCTT CCACCTTTTC  
 3181 CGCGTGTTC GCAGAAACGT GGCTGCCCTG GTTCACCAAC CGGGAAACGG TCTGATAAGA  
 3241 GACACCGGCA TACTCTCGA CATCGTATAA CGTTACTGGT TTCAACATTCA CCACCCCTGAA  
 3301 TTGACTCTCT TCCGGGGCCT ATCATGCCAT ACCCGCGAAG GTTTTGCCTC ATTGATGGT  
 3361 GTCAACGTTAA ATGCCGCTTC GCCTTCGCGC GCGAATTGCA AGCTCTGCAT TAATGAATCG  
 3421 GCCAACGCGC GGGGAGAGGC GGTTTGCCTA TTGGCGCTC TTCCGCTTCC TCGCTCACTG  
 3481 ACTCGCTGCG CTGGTGTGTT CGGCTGCCG GAGCGGTATC AGCTCACTCA AAGGCGGTAA  
 3541 TACGGTTATC CACAGAAATCA GGGGATAAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC  
 3601 AAAAGGCCAG GAACCGTAAA AAGGCCGCTG TGCTGGCGTT TTTCATAGG CTCCGGCCCC  
 3661 CTGACGAGCA TCACAAAAAT CGACGCTAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT  
 3721 AAAGATACCA GGCCTTCCC CCTGGAAGCT CCCTCGTGC CGTCTCTGTT CCGACCCCTGC  
 3781 CGCTTACCGG ATACCTGTCC GCCTTCTCC CTTCGGGAAG CGTGGCGCTT TCTCAATGCT  
 3841 CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGC TGTGTGCACG  
 3901 AACCCCCCGT TCAGCCGAC CGCTGCGCT TATCCGGTAA CTATCGTCTT GAGTCCAACC  
 3961 CGGTAAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA  
 4021 GGTATGTAGG CGGTGCTACA GAGTTCTGAG AGTGGTGGCC TAACTACGGC TACACTAGAA  
 4081 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAA AGAGTTGGTA  
 4141 GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG TTTTTTGTG TGCAAGCAGC  
 4201 AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCCTT GATCTTTCT ACGGGGTCTG  
 4261 ACGCTCAGTG GAACGAAAAC TCACGTTAAAG GGATTTGGT CATGAGATTA TCAAAAGGA  
 4321 TCTTCACCTA GATCTTTTA AATTAAAAAT GAAGTTTAA ATCAATCTAA AGTATATATG  
 4381 AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC TCAGCGATCT  
 4441 GTCTATTTCG TTCACTCCATA GTTGCCTGAC TCCCCGTGTT GTAGATAACT ACGATAACGGG  
 4501 AGGGCTTACC ATCTGGCCCC AGTGCCTGCAA TGATACCGCG AGACCCACGC TCACCCGGCTC  
 4561 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCGA GCGCAGAAAGT GGTCTGCAA  
 4621 CTTTATCCGC CTCCATCCAG TCTATTAAATT GTTGCCTGGGA AGCTAGAGTA AGTAGTTCGC  
 4681 CAGTTAATAG TTGCGCAAC GTTGTGCCA TTGCTACAGG CATCGTGGT TCACCGCTCGT  
 4741 CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC  
 4801 CCATGTTGTG CAAAAAAGCG GTTACGCTCT TCGGTCTCTC GATCGTTGTC AGAAGTAAGT  
 4861 TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC  
 4921 CATCCGTAAG ATGCTTTCT GTGACTGGTG AGTACTCAC CAAAGTCATTC TGAGAATAGT  
 4981 GTATGCGCG ACCGAGTTGC TCTTGGCCCG CGTCAATACG GGATAATACC GCGCCACATA  
 5041 GCAGAACTTT AAAAGTGTCTC ATCATTGGAA AACGTTCTTC GGGGGAAA CTCTCAAGGA  
 5101 TCTTACCGCT GTTACGATCC AGTTCGATGT AACCCACTCG TGCACCCAAAC TGATCTTCAG  
 5161 CATCTTTTAC TTTCACCCAGC GTTCTGGGT GAGCAAAAC AGGAAGGCAA AATGCCGCAA  
 5221 AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCTT TTTCAATATT  
 5281 ATTGAAGCAT TTATCAGGGT TATTGTCATCA TGAGCGGATA CATATTTGAA TGTATTAGA  
 5341 AAAATAAAACA AATAGGGGTT CCGCGCACAT TTCCCGAAA AGTGCACCT GAAATTGTAA  
 5401 ACGTTAATAT TTGTTAAAA TTCGCGTTAA ATTTTGTAA AATCAGCTCA TTTTTTAACC  
 5461 AATAGGCCGA AATCGGAAA ATCCCTTATA AATCAAAAGA ATAGACCGAG ATAGGGTTGA  
 5521 GTGTTGTCTC AGTTTGGAAC AAGAGTCCAC TATTAAGAA CGTGGACTCC AACGTCAAAG  
 5581 GGGAAAAAAC CGTCTATCAG GGCATGGCC CACTACGTGA ACCATCACCC TAATCAAGTT  
 5641 TTTGGGGTC GAGGTGCCGT AAAGCACTAA ATCGGAACCC TAAAGGGAGC CCCCAGTTA  
 5701 GAGCTTGACG GGGAAAGCG GCGAACGTGG CGAGAAAGGA AGGAAAGAAA GCGAAAGGAG  
 5761 CGGGCGCTAG GGCCTGCGA AGTGTAGCGG TCACGCTGCG CGTAACCCAC ACACCCGCCG  
 5821 CGCTTAATGC GCCGCTACAG GGCCTGCTCA TTGCGCCATTG AGGCTGCGCA ACTGTTGGGA  
 5881 AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG GATGTGCTGC  
 5941 AAGGGGATTA AGTTGGG

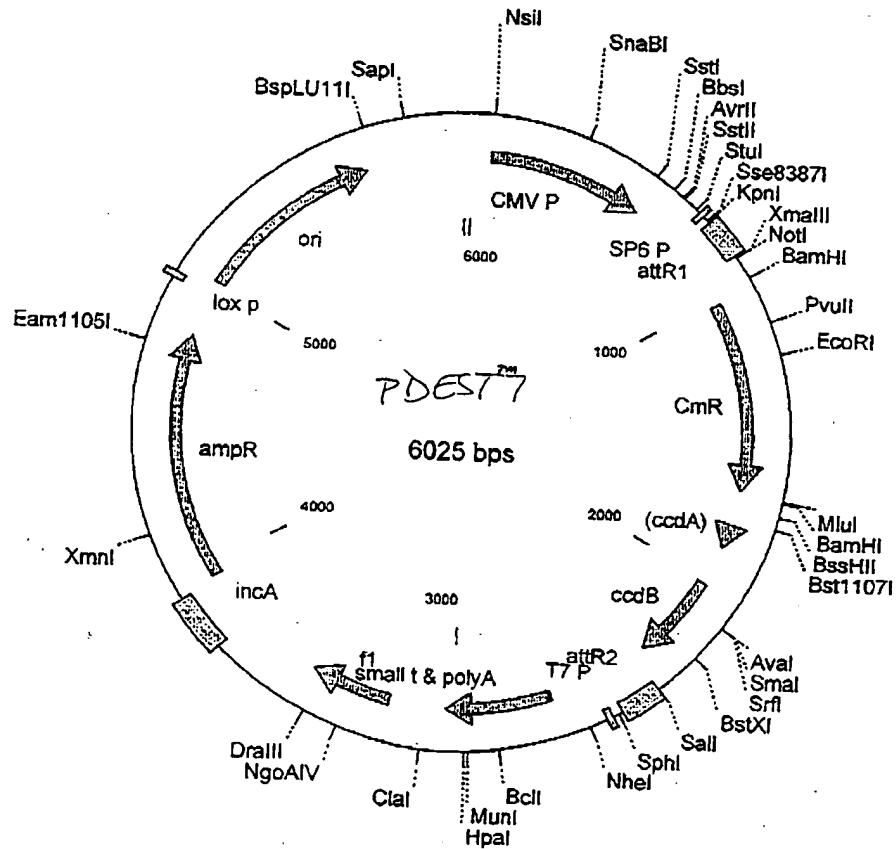
FIGURE 26D

56/240

Figure 27A: PDEST7

## CMV promoter for eukaryotic expression

970 cca ttg acg caa atg ggc ggt agg cgt gta cgg tgg gag gtc tat ata agc  
 ggt aac tgc gtt tac ccc cca tcc gca cat gcc acc ctc cag ata tat tcg  
 1021 aga gct cgt tta gtg aac cgt cag atc gcc tgg aga cgc cat cca cgc tgt  
 tct cga gca aat cac ttg gca gtc tag cgg acc tct gcg gta ggt gcg aca  
 CMV enhancer / promoter  
 1072 ttt gac ctc cat aga aga cac cgg gac cga tcc agc ctc cgg act cta gcc  
 aaa ctg gag gta tct tct gtg gcc ctt gct agg tcc gag gcc tga gat cgg  
 1123 tag gcc gcg gag cgg ata aca att tca cac agg aaa cag cta tga cca cta  
 atc cgg cgc ctc gcc tat tgt taa agt gtg tcc ttt gtc gat act ggt gat  
 1174 ggc ttt tgc aaa aag cta ttt agg tga cac tat aga agg tac gcc tgc tgg  
 ccg aaa acg ttt ttc gat aaa tcc act gtg ata tct tcc atg cgg acg tcc  
 1225 tac EcorI Kpn EcoRI attR1 Pst  
 SspI attR1  
 atg gcc agg cct taa ggg tag tgt tca aac atg ttt tct cga ctc gct tct  
 atg gcc agg cct taa ggg tag tgt tca aac atg ttt tct cga ctc gct tct



57/260

## pDEST7 6025 bp (rotated to position 2800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..589	CMV promoter
906..782	attR1
1015..1674	CmR
1794..1878	inactivated ccdA
2016..2321	ccdB
2362..2486	attR2
2671..3033	small t & polyA
3227..3502	f1
3962..4822	ampR
5022..5661	ori

1 ATTATCATGA CATTAAACCTA TAAAAATAGG CGTAGTACGA GGCCCTTTCA CTCATTAGAT  
 61 GCATGTCGTT ACATAACTTA CCGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG  
 121 CCCATTGACG TCAATAATGA CGTATGTTCC CATACTAACG CCAATAGGGGA CTITTCATTG  
 181 ACGTCAATGG GTGGAGTATT TACCGTAAAC TGCCCACCTTG GCAGTACATC AAGTGTATCA  
 241 TATGCCAAGT ACGCCCCCTA TTGACGTCAA TGACGGTAAA TGGCCCGCCT GGCATTATGC  
 301 CCAGTACATG ACCTTATGGG ACTTTCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC  
 361 TATTACCATG GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC  
 421 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTGTTT GGCACCAAAA  
 481 TCAACGGGAC TTTCCAAAT GTCGTAACAA CTCCGCCCCA TTGACGCAAAG TGGGCGGTAG  
 541 GCGTGTACGG TGGGAGGCT ATATAAGCAG AGCTCGTTA GTGAACCGTC AGATCGCCTG  
 601 GAGACGCCAT CCACGCTGTT TTGACCTCCA TAGAAAGAC CGGGACCGAT CCAGCCTCCG  
 661 GACTCTAGCC TAGGCCGCGG AGCGGATAAC AATTCACAC AGGAAACAGC TATGACCATT  
 721 AGGCCTTTGC AAAAGCTAT TTAGGTGACA CTATAGAAGG TACGCCCTGCA GGTACCGGAT  
 781 CACAAGTTTG TACAAAAAAAG CTGAACGAGA AACGTAAAT GATATAAAATA TCAATATATT  
 841 AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC ATATCCAGTC  
 901 ACTATGGCGG CCGCAITAGG CACCCCAAGGC TTTCACACTT ATGCTTCGGG CTCGTATAAT  
 961 GTGTGGATTG TGAGTGTAGGA TCCGTCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG  
 1021 AAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCCAAT GGCACTGTAAGA ACACATTTT  
 1081 GAGGCATTTC AGTCAGTTG TCAATGTACCA TATAACAGA CGTTCAGCT GGATATTACG  
 1141 GCCTTTTAAAGACCGTAAAG GAAAATAAAG CACAGTTT ATCCGGCCTT TATTACACATT  
 1201 CTTGCCCGCC TGATGAATGC TCATCCGGAA TTCCGTTGCA CAATGAAAAGA CGGTAGCTG  
 1261 GTGATATGGG ATAGTGTTCAC CCCTGTTAC ACCGGTTTCC ATGAGCAAAC TGAAACGTTT  
 1321 TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT ATATTGCAA  
 1381 GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT TGAGAATATG  
 1441 TTTTCGTCT CAGCCAATCC CTGGGTGAGT TTCAACGAGT TTGATTTAA CGTGGCCAAT  
 1501 ATGGACAAC TCTTCGCCCC CGTTTCACC ATGGGCAAAT ATTATACCCA AGGGGACAAG  
 1561 GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCGTCT GTGATGGCTT CCATGTCGGC  
 1621 AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC GTAAACCGGT  
 1681 GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TGCGCGCTGA TTTTTGCGGT  
 1741 ATAAGAATAT ATACTGATAT GTATACCCGA AGTATGTCAA AAAGAGGGTGT GCTATGAAGC  
 1801 AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA TATATGATGT  
 1861 CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAATGAA GCCCCTCGTC TCGTGGCGA  
 1921 ACGCTGGAAA CGGGAAAATC AGGAAGGGAT GGCTGAGGTC GCCCCGGTTA TTGAAATGAA  
 1981 CGGCTTTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT TACACCTATA  
 2041 AAAGAGAGAG CCGTTATCGT CTGTTTGTGG ATGTACAGAG TGATATTATT GACACGCCG  
 2101 GGGGACGGAT GGTGATCCCC CTGGCCAGTG CACGTCTGCT GTCAGATAAA GTCTCCCGT  
 2161 AACTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC ACCGATATGG  
 2221 CCAGTGTGCC GGTCTCCGTT ATCGGGGAAAG AAGTGGCTGA TCTCAGCCAC CGCGAAAATG  
 2281 ACATCAAAA CGCCATTAAAC CTGATGTTCT GGGGAATATA AATGTCAGGC TCCCTTATAC  
 2341 ACAGCCAGTC TGCAAGGTCGA CCATAGTGAC TGGATATGTT GTGTTTACA GTATTATGTA  
 2401 GTCTGTTTT TATGCAAAAT CTAATTTAAT ATATGATAT TTATATCATT TTACGTTCT  
 2461 CGTCAGCTT TCTTGTACAA AGTGGTGCATGC GCGTGCATAG GACGTCATAG CTCTCTCCCT  
 2521 ATAGTGAGTC GTATTATAAG CTAGGCACTG GCGTCTGTT TACAACGTCG TGACTGGAA-

FIGURE 27B

88/240

2581 AACTGCTAGC TTGGGATCTT TGTGAAGGAA CCTTACTTCT CTGGTGTGAC ATAATTGGAC  
 2641 AAACCTACCA CAGAGATTAA AAGCTCTAAG GTAATATAAA AATTTTTAAG TGTATAATGT  
 2701 GTTAAACTAG CTGCATATGC TTGCTGCTTG AGAGTTTGC TTACTGAGTA TGATTATGA  
 2761 AAATATTATA CACAGGAGCT AGTGTACCA ATTGTGTTGTG TATTTAGAT TCACAGTCCC  
 2821 AAGGCTCATT TCAGGCCCCCT CAGTCCTCAC AGTCTGTTCA TGATCATAAT CAGCCATACC  
 2881 ACATTTGTAG AGGTTTACT TGCTTTAAAA AACCTCCCAC ACCTCCCCCT GAACCTGAAA  
 2941 CATAAAATGA ATGCAATTGT TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGTTACAAA  
 3001 TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTT TTTCACTGCA TTCTAGTTGT  
 3061 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCCTGA TCGATCCTGC ATTAATGAAT  
 3121 CGGCCAACGC GCGGGGAGAG GCGGTTTGC GATTGGCTGG CGTAATAGCG AAGAGGCCG  
 3181 CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG  
 3241 CGCGCATTAA AGCGCGGCGG GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG  
 3301 CGCCCTAGCG CCCGCTCCTT TCGCTTCTT CCCTTCCCTT CTCGCCACGT TCGCCGGCTT  
 3361 TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTAGGGTTC CGATTTAGTG CTTTACGGCA  
 3421 CCTCGACCCCC AAAAAGACTG ATTAGGGTGA TGTTCACTG AGTGGGCCAT CGCCCTGATA  
 3481 GACGGTTTT CGCCCTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA  
 3541 AACTGGAACA ACACTCAACC CTATCTCGGT CTAACTCTTT GATTTATAAG GGATTTGCC  
 3601 GATTTGGGCC TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTAA  
 3661 CAAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC  
 3721 TATTGTTTA TTTTCTAAA TACATTCAA TATGTATCCG CTCATGCCAG GTCTGGACT  
 3781 GGTGAGAACG GCTTGTCTGG CAGCTTCGAT GTGTGCTGGA GGGAGAATAA AGGTCTAAGA  
 3841 TGTGCGATAG AGGGAAGTCG CATTGAATT TGTGCTGTG AGGGATCGCT GGTATCAAAT  
 3901 ATGTGTGCC ACCCTGGCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA  
 3961 AAGGAAGAGT ATGAGTATTAC AACATTCCG TGTCGCCCTT ATTCCTTTT TTGCGGCATT  
 4021 TTGCCCTCCT GTTTTGTCT ACCCAGAAC GCTGGTGAAG GTAAAGATG CTGAAGATCA  
 4081 GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTAAC AGCGTAAGA TCCTTGAGAG  
 4141 TTTTCGCCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC  
 4201 GGTATTATCC CGTATTGACG CGGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA  
 4261 GAATGACTTG GTTGGACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT  
 4321 AAGAGAATTAA TGCAGTGTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT  
 4381 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT  
 4441 AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA  
 4501 CACCAAGATG CCTGTAGCAA TGGCAACAAAC GTTGCAGCAA CTATTAACGT GCGAACTACT  
 4561 TACTCTAGCT TCCCGGCAAA AATTAAAGA CTGGATGGAG GCGGATAAAAG TTGCAAGGACC  
 4621 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA  
 4681 GCGTGGGTCT CGCGGTATCA TTGCAAGCACT GGGGCCAGAT GTTAAGCCCT CCCGTATCGT  
 4741 AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAATAGAC AGATCGCTGA  
 4801 GATAAGGTGCC TCACTGATTAA AGCATTGGTA ACTGTCAAGAC CAAGTTTACT CATATATACT  
 4861 TTAGATTGAT TTTAAACTTC ATTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTG  
 4921 TAATCTCATG CCATAACTTC GTATAATGTA TGCTATACGA AGTTATGGCA TGACCAAAAT  
 4981 CCCTTAACGT GAGTTTCTG TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC  
 5041 TTCTTGAGAT CCTTTTTTC TGCGCTTAAT CTGCTGCTTG CAAACAAAAA AACCAACCGCT  
 5101 ACCAGCGGTG GTTGTGTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACTGG  
 5161 CTTCAGCAGA GCGCAGATAAC CAAATACTGT CCTCTAGTG TAGCCGTAGT TAGGCCACCA  
 5221 CTTCAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAAGTGGC  
 5281 TGCTGCCAGT GGCAGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA  
 5341 TAAGGCGCAG CGGTGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC  
 5401 GACTTACACC GAACTGAGAT ACCTACAGCG TGAGCATTGA GAAAGCGCCA CGCTTCCGA  
 5461 AGGGAGAAAG GCGGACAGGT ATCCGTTAAC CGGCAGGGTC GGAACAGGAG AGCGCACGAG  
 5521 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCCT GTGGGTTTC GCCACCTCTG  
 5581 ACTTGAGCGT CGATTTTGTG ATGCTGTC AGGGGGCGG AGCCTATGGA AAAACGCCAG  
 5641 CAACCGGGCC TTTTTACGGT CCTGCGCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC  
 5701 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGC TTTGAGTGAAG CTGATACCGC  
 5761 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCACTGAGC GAGGAAGCGG AAGAGCGCCC  
 5821 AATACGCAAA CCGCCTCTCC CGCGCGGTG GCCGATTCAAT TAATGCAGAG CTTGCAATT  
 5881 GCGCGTTTT CAATATTAA GAAGCAATTAA TCAGGGTTAT TGTCTCATGA GCGGATACAT  
 5941 ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG CGCACATTTC CCCGAAAAGT  
 6001 GCCACCTGAC GTCTAAGAAA CCATT

FIGURE 27C

59/240

**Figure 28A: pDEST8 Polyhedron Promoter, Baculovirus Transfer Plasmid**

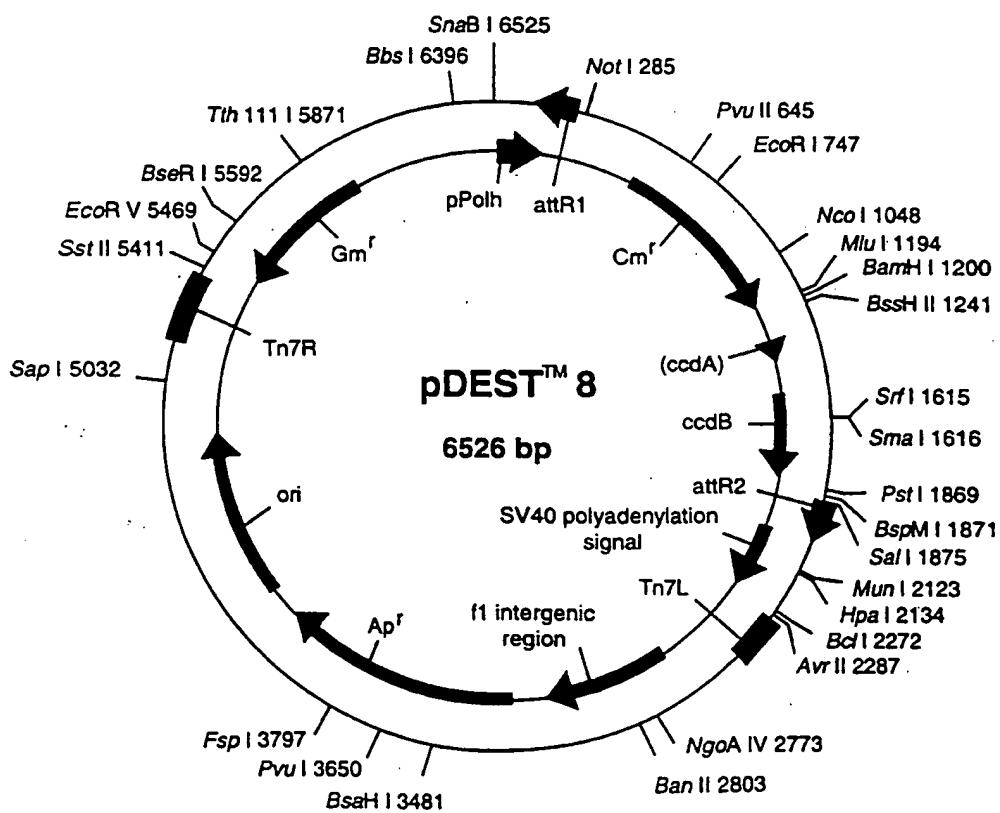
**AccI**

1 cgt ata ctc cgg aat att aat aga tca tgg aga taa tta aaa tga taa cca  
   gca tat gag gcc tta taa tta tct aat acc tct att aat ttt act att ggt

52 tct cgc aaa taa ata agt att tta ctg ttt tcg taa cag ttt tgt aat aaa  
   aga gcg ttt att tat tca taa aat gac aaa agc att gtc aaa aca tta ttt

103 aaa acc tat aaa tat tcc gga tta ttc ata ceg tcc cac cat cgg ggg agg  
   ttt tgg ata ttt ata agg cct aat aag tat ggc agg gtg gta gcc cgc gcc

154 (Bam) attR1 attR2  
   atc atc aca agt tgt tac aaa aa gct gaa cga gaa a g t aa d a t g a t  
   tag tag tgt tca aac atg ttt tgc cga ctt gct ttc tgc att tta ctt tat



60/240

## pDEST8 6526 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
23..152	Ppolh
284..160	attR1
534..1193	CmR
1313..1397	inactivated ccdA
1535..1840	ccdB
1881..2005	attR2
2766..3146	f1
3240..4090	ampR
4289..4869	ori
5564..6496	genR

1 CGTATACTCC GGAATATTAA TAGATCATGG AGATAATTAA AATGATAACC ATCTCGCAA  
 61 TAAATAAGTA TTTTACTGTT TTCGTAACAG TTTGTAATA AAAAACCTA TAAATATTCC  
 121 GGATTAACTCA TACCGTCCC CCATCGGGCG CGGATCATCA CAAGTTGTA CAAAAAAGCT  
 181 GAAACGAGAAA CGTAAATGTA TATAAATATC AATATATTAA ATTAGATTT GCATAAAAAA  
 241 CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC TATGGCGGCC GCTAAGTTGG  
 301 CAGCATCACC CGACGCACCT TGCGCCGAAT AAATACCTGT GACGGAAGAT CACTTCGAG  
 361 AATAAAATAAA TCCTGGTGTG CCTGTTGATA CGGGGAAGCC CTGGGCCAAC TTTTGGCGAA  
 421 AATGAGACGT TGATCGGCAC GTAAGAGGTT CCAACTTCA CCATAATGAA ATAAGATCAC  
 481 TACCGGGCGT ATTTTTTGAG TTATCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA  
 541 AAAAAATCAC TGGATATACC ACCGTTGATA TATCCCAATG GCATCGTAA GAACATTTTG  
 601 AGGCATTTCA GTCACTGTG CAATGTACCT ATAACCAAGAC CGTTCAGCTG GATATTACGG  
 661 CCTTTTAAAG GACCGTAAAG AAAAATAAGC ACAAGTTTA TCCGGCCTTT ATTACACATT  
 721 TTGCCCCGCT GATGAATGCT CATCCGGAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG  
 781 TGATATGGGA TAGTGTTCAC CCTGTTTACA CCGTTTCCA TGAGCAAAC GAAACGTTT  
 841 CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG  
 901 ATGTCGGCTG TTACGGTGAAC CCTGGCCT ATTCCCTAA AGGGTTTATT GAGAATATGT  
 961 TTTTCGTCTC AGCCAATCCC TGGGTGAGTT TCACCAAGTT TGATTTAAC GTGGCCAATA  
 1021 TGGACAACCTT CTTCGCCCCC GTTTTCACCA TGGGCAAATA TTATACGCAA GGCACAAAGG  
 1081 TGCTGATGCC GCTGGCGATT CAGGTTCATC ATGCCGTCTG TGATGGCTTC CATGTCGGCA  
 1141 GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGGGCG TAAACGGCTG  
 1201 GATCCGGCTT ACTAAAAGCC AGATAACAGT ATGCGTATTG GCGCCTGAT TTTGCGGTA  
 1261 TAAGAATATA TACTGATATG TATACCCGAA GTATGTCAA AAGAGGTGTG CTATGAAGCA  
 1321 GCGTATTACA GTGACAGTTG ACAGCAGAC CTATCAGTTG CTCAAGGCAT ATATGATGTC  
 1381 AATATCTCCG GTCTGGTAAG CACAAACATG CAGAATGAAG CCCGTCGCT GCGTGGCAA  
 1441 CGCTGGAAAG CGGAAAATCA GGAAGGATG GCTGAGGTCG CCCGGTTTAT TGAAATGAAC  
 1501 GGCTCTTTTG CTGACGAGAA CAGGGACTGG TGAAATCGAG TTAAAGGTTT ACACCTATAA  
 1561 AAGAGAGAGC CGTTATCGTC GTTTGTGGA TGTACAGAGT GATATTATTG ACACGCCGG  
 1621 GCGACGGATG GTGATCCCCC TGGCACTGC ACGCTGCTG TCAGATAAG TCTCCGTGA  
 1681 ACTTTACCCG GTGGTGCATA TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC  
 1741 CAGTGTGCCG GTCTCCGTTA TCGGGGAAGA AGTGGCTGAT CTCAGCCACC GCGAAAATGA  
 1801 CATAAAAAAC GCCATTAACC TGATGTTCTG GGGAAATATAA ATGTCAGGCT CCCTTATACA  
 1861 CAGCCAGTCT GCAGGTCGAC CATAGTGAATG GGTATGTTG TGTGTTACAG TATTATGTAG  
 1921 TCTGTTTTTATG CAAATTTAATA TATTGATATT TATATCATT TACGTTTCTC  
 1981 GTTCAGCTTT CTTGTACAAA GTGGTGATAG CTTGTCGAGA AGTACTAGAG GATCATAATC  
 2041 AGCCATACCA CATTGTTAGA GGTGTTACTT GCTTTAAAAA ACCTCCCACA CCTCCCCCTG  
 2101 AACCTGAAAC ATAAAATGAA TGCAATTGTT GTTGTAACT TGTGTTATTG AGCTTATAAT  
 2161 GGTTACAAAT AAAGCAATAG CATCACAAAT TTACAAATAA AAGCATTTC TTCACACTGC  
 2221 TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGTCGGAT CTGATCACTG  
 2281 CTTGAGCCTA GGAGATCCGA ACCAGATAAG TGAAATCTAG TTCCAAACTA TTTTGTCTT  
 2341 TTTAATTTC GTATTAGCTT ACGACGCTAC ACCCAGTTCC CATCTATTG GTCACCTTC  
 2401 CCTAAATAAT CCTTAAAAAC TCCATTCCCA CCCCTCCAG TTCCAAACTA TTTTGTCCGC  
 2461 CCACAGCGGG GCATTTCTC TCCCTGTTATG TTTTTAATCA AACATCCCTGC CAACTCCATG  
 2521 TGACAAACCG TCATCTTCGG CTACTTTTC TCTGTCACAG AATGAAAATT TTTCTGTCAT-

FIGURE 28B

61/240

2581 CTCTCGTTA TTAATGTTG TAATTGACTG AATATCAACG CTTATTTGCA GCCTGAATGG  
 2641 CGAATGGACG CGCCCTGTAG CGCGCATTAA AGCGCGGCCG GTGTTGGTGGT TACGCGCAGC  
 2701 GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCTTCCTTT  
 2761 CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC  
 2821 CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT  
 2881 AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTGGA CGTTGGAGTC CACGTTCTTT  
 2941 AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC CTATCTCGGT CTATTCTTT  
 3001 GATTTATAAG GGATTTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT GATTTAACAA  
 3061 AAATTTAACG CGAATTTAA CAAAATATTA ACGTTACAA TTTCAGGTGG CACTTTCCG  
 3121 GGAAATGTGC GCGGAACCCC TATTTGTTA TTTTTCTAAA TACATTCAA TATGTATCCG  
 3181 CTCATGAGAC AATAACCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT  
 3241 ATTCAACATT TCCGTGTGCG CCTTATTCCTT TTTCGCGG CATTTCGCT TCCTGTTTT  
 3301 GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACCGAGTG  
 3361 GTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTCG CCCCAGAAGAA  
 3421 CGTTTCCAA TGATGAGCAC TTTAAAGTT CTGCTATGTC GCGCGGTATT ATCCCGTATT  
 3481 GACGCCGGGC AAGAGCAACT CGGTGCCCGC ATACACTATT CTCAAAATGA CTTGGTTGAG  
 3541 TACTCACCAG TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT  
 3601 GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACCTAC TTCTGACAAC GATCGGAGGA  
 3661 CCGAAGGAGC TAACCGCTT TTGACAAAC ATGGGGGATC ATGTAACCTG CCTTGATCGT  
 3721 TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCCTGTA  
 3781 GCAATGGAA CAAACGTTGGC CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCG  
 3841 CAACAATTAA TAGACTGGAT GGAGGGGGAT AAAGTTGAG GACCACTTCT GCGCTCGGCC  
 3901 CTTCCGGCTG GCTGGTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT  
 3961 ATCATTGAG CACTGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT CTACACGACG  
 4021 GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG  
 4081 ATTAAGCATT CGTAACCTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA  
 4141 CTTCACTTTT AATTTAAAG GATCTAGGTG AAGATCCTT TTGATAATCT CATGACCAAA  
 4201 ATCCCTTAAC GTGAGTTTC GTTCCACTGA GCGTCAGACCC CGTAGAAAA GATCAAAGGA  
 4261 TCTTCTTGAG ATCCTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCACCG  
 4321 CTACCAAGCGG TGGTTTGTGTT GCCGGATCAA GAGCTACCAA CTCTTTTCC GAAGGTAACT  
 4381 GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCTTCTAG TGTAGCCGTA GTTAGGCCAC  
 4441 CACTTCAAGA ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAAGTG  
 4501 GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTGG ACTCAAGACG ATAGTTACCG  
 4561 GATAAGGCAGC AGCGGTGGGG CTGAACGGGG GGTTCGTGCA CACAGCCAG CTTGGAGCGA  
 4621 ACGACCTACA CGAAGCTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC  
 4681 GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG  
 4741 AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTATAGTC CTGTCGGGTT TCGCCACCTC  
 4801 TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGG GGAGCCTATG GAAAACGCC  
 4861 AGCAACCGGG CCTTTTACG GTTCTGGCC TTTTGCTGG CTTTTGCTCA CATGTTCTTT  
 4921 CCTGCGTTAT CCCCTGATTC TGTTGATAAC CGTATTACCG CTTTGAGTG AGCTGATACC  
 4981 GCTGCCCGCA GCGGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC  
 5041 CTGATGGGGT ATTTCTCCT TACCGATCTG TGCGGTATTT CACACCGCAG ACCAGCCCG  
 5101 TAACCTGGCA AAATCGGTTA CGGTTGAGTA ATAAATGGAT GCCCTCGTA AGCGGGTGTG  
 5161 GCGGACAAT AAAGTCTTAA ACTGAACAAA ATAGATCTAA ACTATGACAA TAAAGTCTTA  
 5221 AACTAGACAG AATAGTTGTA AACTGAAATC AGTCCAGTTA TGCTGTGAAA AAGCATAACTG  
 5281 GACTTTGTT ATGGCTAAG CAAACTCTTC ATTTCTGAA GTGCAAATTG CCCGCTGTAT  
 5341 TAAAGAGGGG CGTGGCCAAG GGCATGGTAA AGACTATATT CGCGGCCCTTG TGACAATTAA  
 5401 CGAACAACT CGCGGCCCGG GAAGGCCGATC TCGGCTTGAA CGAATTGTTA GGTGGCGGTA  
 5461 CTTGGGTCGA TATCAAAGTG CATCACTTCT TCCCGTATG CCAACTTTGT ATAGAGAGCC  
 5521 ACTGCCGGAT CGTCACCGTA ATCTGCTTGC ACGTAGATCA CATAAGCACC AAGCGCGTTG  
 5581 CCCTCATGCT TGAGGAGATT GATGAGCGCG GTGGCAATGC CCTGCCTCCG GTGCTCGCCG  
 5641 GAGACTGCCA GATCATAGAT ATAGATCTCA CTACGCGGCT GCTCAAACCT GGGCAGAACG  
 5701 TAAGCCCGA GAGCGCCAAC AACCGCTTCT TGGTCGAAGG CAGCAAGCGC GATGAATGTC  
 5761 TTACTACGGA GCAAGTCCCC GAGGTAATCG GAGTCCGGCT GATGTTGGGA GTAGGTGGCT  
 5821 ACGTCTCCGA ACTCACCGACC GAAAAGATCA AGAGCAGGCC GCATGGATTT GACTTGGTCA  
 5881 GGGCGGAGCC TACATGTGCG AATGATGCCC ATACTTGAGC CACCTAACTT TGTTTGTAGGG  
 5941 CGACTGCCCT GCTGCGTAAC ATCGTTGCTG CTGCGTAACA TCGTTGCTGC TCCATAACAT  
 6001 CAAACATCGA CCCACGGCGT AACCGCCTTG CTGCTTGGAT GCCCGAGGCA TAGACTGTAC

FIGURE 28C

62/240

6061 AAAAAAACAG TCATAACAAG CCATGAAAAC CGCCACTGCG CCGTTACCAC CGCTGCCTTC  
6121 GGTCAGGTG CTGGACCACT TGCAGTGCAGC CATACGCTAC TTGCATTACA GTTTACGAAC  
6181 CGAACAGGCT TATGTCAACT GGGTTCTGTC CTTCATCCGT TTCCACGGTG TGCGTCACCC  
6241 GGCAACCTTG GGCAGCAGCG AAGTCGAGGC ATTCTGTCC TGGCTGGCGA ACGAGCGCAA  
6301 GGTTTCGGTC TCCACGCATC GTCAGGCATT GGCGGCCTTG CTGTTCTTCT AC GGCAAGGT  
6361 GCTGTGCACG GATCTGCCCT GGCTTCAGGA GATCGGAAGA CCTCGGCCGT CGCGCGCTT  
6421 GCCGGTGGTG CTGACCCCCGG ATGAAGTGGT TCGCATCCTC GGTTTCTGG AAGGCGAGCA  
6481 TCGTTGTTC GCCCAGGACT CTAGCTATAG TTCTAGTGGT TGGCTA

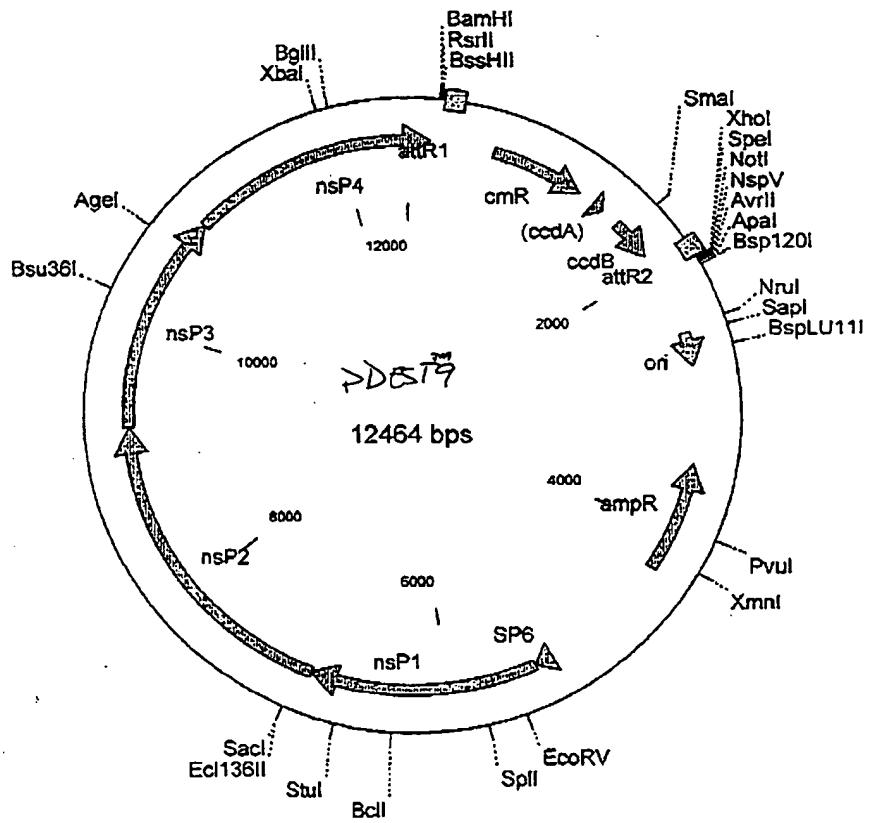
FIGURE 28D

63/240

Figure 29A: pDEST9

## Semliki Forest Virus vector

103 ttg gcg agg gac att aag ggc ttt aag aaa ttg aga gga cct gtt ata ~~cac~~  
 aac cgc tcc ctg taa ttc cgc aaa ttc ttt aac tct cct gga caa tat ~~gtg~~  
 154 ~~cgc tac ggc ggt cct aca ttg gtg cgt taa tac aca gaa ttc tga ttg gat~~ Bam  
~~gag atg cgg cca gga tct aac ccc gca att atg tgt ctt aag act aac cta~~  
 205 ~~ccc ggt ccc aag cgc gct ttc cca tca [aca agt ttg tac aaa aad oct gpa~~ <sup>Bam</sup>  
~~ggg cca ggc ttc gcg cga aag ggt agt tgt tca aac atg ttt tcc cga tct~~ <sup>Sst</sup> <sup>attR1</sup>



64/240

## pDEST9 12464 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
355..232	attR1
605..1264	CmR
1384..1468	inactivated ccdA
1606..1911	ccdB
1952..2078	attR2
2532..2782	ori
3482..4282	ampR
5232..5365	SP6 promoter
5365..6965	nsP1:non-structural protein 1
6965..9265	nsP2:non-structural protein 2
9265..10865	nsP3:non-structural protein 3
10865..161	nsP4:non-structural protein 4

1 AGCAAGTGGT TCCGGACAGG CTTGGGGGCC GAACTGGAGG TGGCACTAAC ATCTAGGTAT  
 61 GAGGTAGAGG GCTGCAAAG TATCCTCATA GCCATGGCCA CCTTGGCGAG GGACATTAAAG  
 121 GCGTTTAAGA AATTGAGAGG ACCTGTTATA CACCTCTACG GCGGTCTTAG ATTGGTGCCT  
 181 TAATACACAG AATTCTGATT GGATCCCGGT CGAAGCGCG CTTTCCCCTAC ACAAGTTTGT  
 241 ACAAAAAAGC TGAACGAGAA ACGTAAATG ATATAAATAT CAATATATTA AATTAGATTT  
 301 TGCATAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGGC  
 361 CGCTAAGTTG GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG TGACGGAAGA  
 421 TCACTTCGCA GAATAAATAA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC CCTGGCCAA  
 481 CTTTGGCGA AAATGAGACG TTGATCGGCA CGTAAGAGGT TCCAACCTTC ACCATAATGA  
 541 AATAAGATCA CTACCGGGCG TATTTTTGTA GTTATCGAGA TTTCAGGAG CTAAGGAAGC  
 601 TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCCAAT GGCATCGTAA  
 661 AGAACATTTC GAGGCATTT AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAAGCT  
 721 GGATATTACG GCCTTTTAA AGACCGTAA GAAAAATAAG CACAAGTTT ATCCGGCCTT  
 781 TATTCACATT CTTGCCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA  
 841 CGGTGAGCTG GTGATATGGG ATAGTGTTCAC CCGTTGTTAC ACCGTTTTC ACCGAGCAAC  
 901 TGAAACGTTT TCATCGCTCT GGAGTGAAATA CCACGACGAT TTCCGGCAGT TTCTACACAT  
 961 ATATTGCAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT  
 1021 TGAGAAATATG TTTTTCGTCAGCCAATCC CTGGGTGAGT TTCACCAGTT TTGATTAA  
 1081 CGTGGCCAAT ATGGACAAC TCTTCGCCCC CGTTTCACC ATGGGCAAAT ATTATACGCA  
 1141 AGGGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT  
 1201 CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC  
 1261 GTAAAGATCT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TCGCGCCTGA  
 1321 TTTTGGCGGT ATAAGAAATAT ATACTGATAT GTATACCCGA AGTATGTCAA AAAGAGGTGT  
 1381 GCTATGAAAG AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA  
 1441 TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAATGAA GCGCGTCGTC  
 1501 TGCCTGCCGA ACGCTGGAAA GCGGAAAATC AGGAAGGGAT GGCTGAGGTC GCGCGGTTA  
 1561 TTGAAATGAA CGGCTTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT  
 1621 TACACCTATA AAAGAGAGAG CCGTTATCGT CTGTTTGTGG ATGTACAGAG TGATATTATT  
 1681 GACACGCCCG GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGCTCGCT GTCAGATAAA  
 1741 GTCTCCCGTG AACTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC  
 1801 ACCGATATGG CCAGTGTGCC GGTCTCCGTT ATCGGGGAAG AAGTGGCTGA TCTCAGCCAC  
 1861 CGCGAAAATG ACATCAAAA CGCCATTAAC CTGATGTTCT GGGGAATATA AATGTCAGGC  
 1921 TCCCTTATAC ACAGCCAGTC TGCAGGTGCA CCATAGTGCAC TGGATATGTT GTGTTTACA  
 1981 GTATTATGTA GTCTGTTTT TATGAAAAG TGCTAATTAA ATATATTGAT ATTTATATCA  
 2041 TTTTACGTTT CTCGTTCAGC TTTCTTGTAC AAAGTGGTGA TGGGAACCTCG AGTTCACTAG  
 2101 TCGATCCCGC GGCGCTTTG GAACCTAGGC AAGCATGCGG GCGCAGTGGG TAATTAATTG  
 2161 AATTACATCC CTACGCAAAC GTTTACGGC CGCCGGTGGC GCGCGCCCGC GGCGGCCCGT  
 2221 CCTTGGCGGT TGCAGGCCAC TCCGGTGGCT CCCGTCGTC CCGACTTCCA GGCCAGCAG  
 2281 ATGCAGCAAC TCATCAGCGC CGTAAATGCG CTGACAATGA GACAGAACGC AATTGCTCCT  
 2341 GCTAGGAGCT TAATTCGACG AATAATTGGA TTTTATTAAATTTGCAAT TGGTTTTAA  
 2401 TATTCCAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA-

FIGURE 29B

65/240

2461 AAAAAAAA AAAAAACTA GAAATCGCGA TTTCTAGTCT GCATTAATGA ATCGGCCAAC  
 2521 GCGCGGGGAG AGGC GGTTTG CGTATTGGC GCTCTTCGC TTCCCTCGCTC ACTGACTCGC  
 2581 TGCGCTCGGT CGTTCGGCTG CGCGAGCGG TATCAGCTCA CTCAAAGGCG STAATACGGT  
 2641 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG  
 2701 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTCCTCCA TAGGCTCCGC CCCCCGTGACG  
 2761 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAAGAT  
 2821 ACCAGGC GTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCCCTTA  
 2881 CCGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCGCGCT  
 2941 GTAGGTATCT CAGTTGGTG TAGGTCGTTG GCTCCAAGCT GGGCTGTGTG CACGAACCCC  
 3001 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAATATCG TCTTGAGTCC AACCCGGTAA  
 3061 GACAGGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG  
 3121 TAGGCGGTGC TACAGAGTTT TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG  
 3181 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT  
 3241 GATCCGGCAA ACAAAACACC GCTGGTAGCG GTGGTTTTTG TGTTTGCAAG CAGCAGATT  
 3301 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC  
 3361 AGTGGAAACGA AAACCTACGT TAAGGGATT TGTTCATGAG ATTATCAAAA AGGATCTTCA  
 3421 CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA  
 3481 CTTGGTCTGA CAGTTACCA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGCTAT  
 3541 TTCGTTCATC CATAGTTGCC TGACTCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT  
 3601 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT  
 3661 TATCAGCAAT AAACAGGCC GCGGAAGGG CCGAGCGCAG AAGTGGTCC GCAACTTTAT  
 3721 CGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT TCGCCAGTTA  
 3781 ATAGTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTACCGC TCGTCGTTTG  
 3841 GTATGGCTTC ATTCA GCTCC GATCAAGGCG AGTTACATGA TCCCCCATGT  
 3901 TGTGCAAAA AGCGGTTAGC TCCCTCGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG  
 3961 CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATT TCTTACTGTC ATGCCATCCG  
 4021 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC  
 4081 GGCGACCGAG TTGCTCTTGC CCGCGTCAA TACGGATAA TACCGCGCC CATAGCAGAA  
 4141 CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTGGGGCG AAAACTCTCA AGGATCTTAC  
 4201 CGCTGTTGAG ATCCAGTTG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT  
 4261 TTACTTTCAC CAGC GTTCTG GGGTGAGCAA AAACAGGAAG GCAAATGCC GCAAAAAGG  
 4321 GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA  
 4381 GCATTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA  
 4441 AACAAATAGG GGTTCGGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA  
 4501 TTATTATCAT GACATTAACT TATAAAATA GGCATATCAC GAGGCCCTT CGTCTCGCGC  
 4561 GTTTGGTGA TGACGGTGA AACCTCTGAC ACATGCA GCTCCAGCT CCCGGAGACG GTCACAGCTT  
 4621 CTGCTAAGC GGATGCCGG AGCAGACAAG CCCGTCAGGG CGCCTCAGCG GGTGTTGGCG  
 4681 GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGGCA GAGT TGACTGAGA GTGCACCCATA  
 4741 TCGACGCTCT CCTTATGCG ACTCCTGCAT TAGGAAGCAG CCCAGTACTA GGTGAGGCC  
 4801 GTTGAGCACC GCGCCGCAA GGAATGGTGC ATGCAAGGAG ATGGCCCCC ACAGTCCCC  
 4861 GGCCACGGGG CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG  
 4921 AGCCGATCT TCCCCATCG TGATGTCGGC GATATAGGCG CCAGCAACCG CACCTGTGGC  
 4981 GCGGGTGATG CGGGCCACGA TGCGTCCGGC GTAGAGGATC TGGCTAGCGA TGACCCGTGCT  
 5041 GATTGGTTCG CTGACCATTT CGGGGGTGC GAAAGCGTTT ACCAGAAACT CAGAAGGTTT  
 5101 GTCCAACCAA ACCGACTCTG ACGGCAGTTT ACGAGAGAGA TGATAGGGTC TGCTTCAGTA  
 5161 AGCCAGATGC TACACAATTG GGCTTGACA TATTGTCGTT AGAACGCGGC TACAATTAA  
 5221 ACATAACCTT ATGATATCATA CACATACGAT TTAGGTGACA CTATAGATGG CGGATGTGTG  
 5281 ACATACACGA CGCCAAAAGA TTTTGTCCA GCTCTCGCCA CCTCCGCTAC GCGAGAGATT  
 5341 AACCAACCCAC GATGGCCGCC AAAGTGCATG TTGATATTGA GGCTGACAGC CCATTCTATCA  
 5401 AGTCTTGCA GAAGGCATTT CGTCGTTG AGGTGGAGTC ATTGCAAGTC ACACCAAATG  
 5461 ACCATGCAA TGCCAGAGCA TTTTCGCA TGGCTACCAA ATTGATCGAG CAGGAGACTG  
 5521 ACAAAAGACAC ACTCATCTTG GATATGGCA GTGCGCTTC CAGGAGAATG ATGTCTACGC  
 5581 CAAATAACCA CTGCGTATGC CCTATGCGCA GCGCAGAAGA CCCCAGAAAGG CTCGATAGCT  
 5641 ACGCAAAAGAA ACTGGCAGCG GCGTCCGGG AGGTGCTGG TAGAGAGATC GCAGGAAAAA  
 5701 TCACCGACCT GCAAGACCGTC ATGGCTACGC CAGACGCTGA ATCTCTAAC TTTTGCTG  
 5761 ATACAGACGT CACGTGTCGT ACGGCAGCCG AAGTGGCCGT ATACCAGGAC GTGTATGCTG  
 5821 TACATGCACC AACATCGCTG TACCATCAGG CGATGAAAGG TGTCAGAACG GCGTATTGGA  
 5881 TTGGGTTTGA CACCACCCCG TTTATGTTG ACGCGCTAGC AGGCAGCTAT CCAACCTACG-

FIGURE 29C

66/240

5941 CCACAAACTG GGCGGACGAG CAGGTGTTAC AGGCCAGGAA CATAGGACTG TGTGCAGCAT  
 6001 CCTTGACTGA GGGAAAGACTC GGCAAACGTG CCATTCTCCG CAAGAAGCAA TTGAAACCTT  
 6061 GCGACACAGT CATGTTCTCG GTAGGATCTA CATTGTACAC TGAGAGCAGA AAGCTACTGA  
 6121 GGAGCTGGCA CTTACCCCTCC GTATTCCACC TGAAAGGTA ACAATCCTTT ACCTGTAGGT  
 6181 GCGATACCAT CGTATCATGT GAAGGGTAGC TAGTTAAGAA AATCACTATG TGCCCCGGCC  
 6241 TGTACCGTAA AACGGTAGGG TACGCCGTGA CGTATCACGC GGAGGGATTC CTAGTGTGCA  
 6301 AGACCACAGA CACTGTAAA GGAGAAAGAG TCTCATTCCC TGTATGCACC TACGTCCCT  
 6361 CAACCATCTG TGATCAAATG ACTGGCATAC TAGGGACCGA CGTCACACCG GAGGACGCAC  
 6421 AGAAGTTGTT AGTGGGATTG AATCAGAGGA TAGTTGTGAA CGGAAGAACAA CAGCGAAACA  
 6481 CTAACACGAT GAAGAACTAT CTGCTTCCGA TTGTTGGCCGT CGCATTAGC AAGTGGCGA  
 6541 GGGAAATACAA GGCAAGACCTT GATGATGAAA AACCTCTGGG TGTCCGAGAG AGGTCACTTA  
 6601 CTTGCTGCTG CTTGTGGGCA TTAAAACGA GGAAGATGCA CACCATGTAC AAGAAACCAAG  
 6661 ACACCCAGAC AATAGTGAAG GTGCCCTCAG AGTTAACTC GTTGTTCATC CCGAGCTAT  
 6721 GGTCTACAGG CCTCGCAATC CCAGTCAGAT CACGCATTAA GATGCTTTTG GCCAAGAAGA  
 6781 CCAAGCGAGA GTTAATACCT GTTCTCGACG CGTCGTCAGC CAGGGATGCT GAACAAAGAGG  
 6841 AGAAGGAGAG GTTGGAGGCC GAGCTGACTA GAGAAGCCTT ACCACCCCTC GTCCCCATCG  
 6901 CGCCGGCGGA GACGGGAGTC GTGACGTCG ACGTTGAAGA ACTAGAGTAT CACGCAGGTG  
 6961 CAGGGTCTGT GGAAACACCT CGCAGCGCGT TGAAAGTCAC CGCACAGCGC AACGACGTAC  
 7021 TACTAGGAAA TTACGTAGTT CTGCTCCCGC AGACCGTGCT CAAGAGCTCC AAGTTGGCCC  
 7081 CCGTGCACCC TCTAGCAGAG CAGGTAAAA TAATAACACA TAACGGGAGG GCCGGCGGTT  
 7141 ACCAGGTCGA CGGATATGAC GGCAGGGTCC TACTACCATG TGGATCGGCC ATTCCGGTCC  
 7201 CTGAGTTTCA GGCTTTGAGC GAGAGCGCCA CTATGGTGTAA CAACGAAAGG GAGTTGTCA  
 7261 ACAGGAAACT ATACCATATT GCGCTTCACG GACCCCTCGCT GAACACCGAC GAGGAGAACT  
 7321 ACGAGAAAGT CAGAGCTGAA AGAACTGACG CCGAGTACGT TTTCGACGTA GATAAAAAAAT  
 7381 GCTGCGTCAA GAGAGAGGAA GCGTCGGGTT TGGTGTGGT GGGAGAGCTA ACCAACCCCC  
 7441 CGTTCCATGA ATTGCGCTAC GAAGGGCTGA AGATCAGGCC GTCGGCACCA TATAAGACTA  
 7501 CAGTAGTAGG AGTCTTTGGG GTTCCGGGAT CAGGCAAGTC TGCTATTATT AAGAGCCTCG  
 7561 TGACCAAACA CGATCTGGTC ACCAGCGCA AGAAGGAGAA CTGCCAGGAA ATAGTTAACG  
 7621 ACGTGAAGAA GCACCGCGGG AAGGGGACAA GTAGGGAAAA CAGTGAECTC ATCCTGCTAA  
 7681 ACGGGTGTG TCGTGCCTGTG GACATCCTAT ATGTGGACGA GGCTTTCGCT TGCCATTCCG  
 7741 GТАCTCTGCT GGCCCTAATT GCTCTTGTAA AACCTCGGAG CAAAGTGGTG TTATGCGGAG  
 7801 ACCCCAAAGCA ATGCGGATTG TTCAATATGA TGCAGCTTAA GGTGAACTTC AACCCAAACA  
 7861 TCTGCACTGA AGTATGTCA AAAAGTATAT CCAGACGTTG CACGCGTCCA GTCACGGCCA  
 7921 TCGTGTCTAC GTTGCACTAC GGAGGCAAGA TGCGCACGAC CAACCCGTGC AACAAACCCA  
 7981 TAATCATAGA CACCAACAGGA CAGACCAAGC CCAAGCCAGG AGACATCGTG TTAACATGCT  
 8041 TCCGAGGCTG GGCAGGACAG CTGAGTTGG ACTACCGTGG ACACGAAGTC ATGACAGCAG  
 8101 CAGCATCTCA GGGCCTCACC CGCAAAGGGG TATACGCCGT AAGGCAGAAG GTGAAT3AAA  
 8161 ATCCCTTGTA TGCCCTGCG TCGGAGCACG TGAATGTACT GCTGACGCGC ACTGAGGATA  
 8221 GGCTGGTGTG GAAAACGCTG GCGGGCGATC CCTGGATTAA GTCTCTATCA AACATTCCAC  
 8281 AGGGTAACCT TACGGCCACA TTGGAAGAAT GGCAAGAAGA ACACGACAAA ATAATGAAGG  
 8341 TGATTGAAGG ACCGGCTGCG CCTGTGGACG CGTTCAGGAA CAAAGCGAAC GTGTGTGG  
 8401 CGAAAAGCCT GGTGCCGTGTC CTGGACACTG CCGGAATCAG ATTGACAGCA GAGGAGTGG  
 8461 GCACCATATAAT TACAGCATTT AAGGAGGACA GAGCTTACTC TCCAGTGGTG GCCTTGATG  
 8521 AAATTGTCAC CAAGTACTAT GGAGTTGACC TGGACAGTGG CTCGTTTCT GCCCCGAAAGG  
 8581 TGTCCCTGTA TTACGAGAAC AACCACGGG ATAACAGACC TGGTGGAAAGG ATGTAT3GAT  
 8641 TCAATGCCGC AACAGCTGCC AGGCTGGAAAG CTAGACATAC CTTCCGTAAAG GGGCAGTGGC  
 8701 ATACGGGCAA GCAGGCAGTT ATCGCAGAAA GAAAATCCA ACCGCTTTCT GTGCTGGACA  
 8761 ATGTAATTCC TATCAACCGC AGGCTGCCGC ACAGCCCTGGT GGCTGAGTAC AAGACGGTTA  
 8821 AAGGCAGTAG GGGTGGAGTGG CTGGTCAATA AAGTAAGAGG GTACCACGTC CTGCTGGTGA  
 8881 GTGAGTACAA CCTGGCTTTG CCTCGACGCA GGGTCACTTG GTTGTACCC CTGAATGTCA  
 8941 CAGGGCGCGA TAGGTGCTAC GACCTAAGTT TAGGACTGCC GGCTGACGCC GGCAGGTTCG  
 9001 ACTTGGTCTT TGTGAACATT CACACGGAAAT TCAGAATCCA CCACTACCGAG CAGTGTGTCG  
 9061 ACCACGCCAT GAAGCTGCAG ATGCTTGGGG GAGATGCGCT ACGACTGCTA AAACCCGGCG  
 9121 GCATCTTGAT GAGAGCTTAC GGATACGGCG ATAAAATCAG CGAAGCCGTT GTTCCCTCCT  
 9181 TAAGCAGAAA GTTCTCGTCT GCAAGAGTGT TGGCCCGGA TTGTGTACCC AGCAATACAG  
 9241 AAGTGTCTT GCTGTTCTCC AACTTTGACA ACAGGAAAGAG ACCCTCTACG CTACACCAGA  
 9301 TGAATACCAA GCTGAGTGCCT GTGTATGCCG GAGAAGCCAT GCACACGGCC GGGTGTGAC  
 9361 CATCCTACAG AGTTAAGAGA GCAGACATAG CCACGTGCAC AGAAGCGGCT GTGGTTAACG-

FIGURE 29d

67/240

9421 CAGCTAACGC CCGTGGAACT GTAGGGGATG GCGTATGCAG GGCGTGGCG AAGAAATGGC  
 9481 CGTCAGCCCT TAAGGGAGCA GCAACACCAG TGGGCACAAT TAAAACAGTC ATGTGCGGCT  
 9541 CGTACCCCGT CATCCACGCT GTAGGCCTA ATTCTCTGC CACGACTGAA GCAGAAGGG  
 9601 ACCGCGAATT GGGCGCTGTC TACCGGGCAG TGGCCGCCGA AGTAAACAGA CTGTCACTGA  
 9661 GCAGCGTAGC CATCCCGCTG CTGTCCACAG GAGTGTTCAG CGCGGAAGA GATAGGCTGC  
 9721 AGCAATCCCT CAACCATCTA TTACACAGCA TGGACGCCAC GGACGCTGAC GTGACCATCT  
 9781 ACTGCAGAGA CAAAAGTGG GAGAAGAAAA TCCAGGAAGC CATTGACATG AGGACGGCTG  
 9841 TGGAGTTGCT CAATGATGAC GTGGAGCTGA CCACAGACTT GGTGAGAGTG CACCCGGACA  
 9901 GCAGCCTGGT GGGTGTAG GGCTACAGTA CCACTGACGG GTGCTGTAC TCGTACTTTG  
 9961 AAGGTACGAA ATTCAACAG GCTGCTATIG ATATGGCAGA GATACTGACG TTGTGGCCA  
 10021 GACTGCAAGA GGCAAACGAA CAGATATGCC TATACCGCCT GGGCGAAACA ATGGACAAACA  
 10081 TCAGATCCAA ATGTCGGGT AACGATTCCG ATTCACTAAC ACCTCCCAGG ACAGTGCCCT  
 10141 GCCTGTGCCG CTACGCAATG ACAGCAGAAC GGATCGCCCG CCTTAGGTCA CACCAAGTTA  
 10201 AAAGCATGGT GGTITGCTCA TCTTTCCCC TCCCGAAATA CCATGTAGAT GGGGTGCAGA  
 10261 AGGTAAAGTG CGAGAAGGTT CTCCCTGTCG ACCCGACGGT ACCTTCAGTG GTTACGCCGC  
 10321 GGAAGTATGC CGCATCTACG ACGGACCACT CAGATCGGTC GTTACGAGGG TTTGACTTGG  
 10381 ACTGGACCAC CGACTCGTCT TCCACTGCCA GCGATACCAT GTGCTACCC AGTTTGCAGT  
 10441 CGTGTGACAT CGACTCGATC TAGCAGCCAA TGGCTCCCAT AGTACTGACG GCTGACGTAC  
 10501 ACCCTGAACC CGCAGGCATC GCGGACCTGG CGGCAGATGT GCACCCCTGAA CCCGCAGACC  
 10561 ATGTGGACCT GGAGAACCCG ATTCCTCCAC CGCGCCCGAA GAGAGCTGCA TACCTTGCCT  
 10621 CCCGCGCGC GGAGCGACCG GTGCCGGCG CGAGAAAGCC GACGCCCTGCC CCAAGGACTG  
 10681 CGTTTAGGAA CAAGCTGCC TTGACGTTCG GCGACTTTGA CGAGCACGAG GTCGATGCGT  
 10741 TGGCCTCCGG GATTACTTTC GGAGACTTCG ACGACGTCCT GCGACTAGGG CGCGCGGGTG  
 10801 CATATATTTT CTCTCTGGAC ACTGGCAGCG GACATTTACA ACAAAATCC GTTACGGCAGC  
 10861 ACAATCTCCA GTGCGCACAA CTGGATGCCG TCCAGGAGGA GAAAATGTAC CCGCCAAAAT  
 10921 TGGATACTGA GAGGGAGAAG CTGTTGCTGC TGAAATGCA GATGCACCCA TCGGAGGCTA  
 10981 ATAAGAGTCG ATACCACTCT CGCAAAGTGG AGAACATGAA AGCCACGGTG GTGGACAGGC  
 11041 TCACATCGGG GGCCAGATTG TACACGGAG CGGACGTAGG CCGCATACCA ACATACGCGG  
 11101 TTCGGTACCC CCGCCCCGTG TACTCCCTA CCGTGTACGA AAGATTCTCA AGCCCCGATG  
 11161 TAGCAATCGC AGCGTGCAAC GAATACCTAT CCAGAAATTAA CCCAACAGTG GCGTCGTAC  
 11221 AGATAACAGA TGAATACGAC GCATACTTGG ACATGGTTGA CGGGTCGGAT AGTTGCTTGG  
 11281 ACAGAGCGAC ATTCTGCCCG GCGAAGCTCC GGTGCTACCC GAAACATCAT GCGTACCA  
 11341 AGCCGACTGT ACGCAGTGCC GTCCCGTCAC CCTTTCAGAA CACACTACAG AACGTGCTAG  
 11401 CGGCTGCCAC CAAGAGAAC TGCAACGTCA CGCAAATGCG AGAAACTACCC ACCATGGACT  
 11461 CGGCAGTGT CAACGTGGAG TGCTTCAAGC GCTATGCCG CTCCGGAGAA TATTGGGAAG  
 11521 AATATGCTAA ACAACCTATC CGGATAACCA CTGAGAACAT CACTACCTAT GTGACCAAAT  
 11581 TGAAAGGCC GAAAGCTGCT GCCTTGTTCG CTAAGACCCA CAACTTGGTT CCGCTGCCAGG  
 11641 AGGTCCCCAT GGACAGATTG ACGGTCGACA TGAAACGAGA TGTCAAAGTC ACTCCAGGG  
 11701 CGAAACACAC AGAGGAAAGA CCCAAAGTCC AGGTAAATTCA AGCAGCGGAG CCATTGGCGA  
 11761 CGCGCTTACCT GTGCGGCATC CACAGGGAT TAGTAAGGAG ACTAAATGCT GTGTTACGCC  
 11821 CTAACGTGCA CACATTGTTT GATATGTCGG CCGAAGACTT TGACGCGATC ATCGCCTCTC  
 11881 ACTTCCACCC AGGAGACCCG GTTCTAGAGA CGGACATTGC ATCATTGAC AAAAGCCAGG  
 11941 ACGACTCCCT GGCTCTTACA GGTTTAATGA TCCTCGAAGA TCTAGGGGTG GATCAGTACC  
 12001 TGCTGGACTT GATCGAGGCA GCCTTGGGG AAATATCCAG CTGTCACCTA CCAACTGGCA  
 12061 CGCGCTTCAAA GTTCGGAGCT ATGATGAAAT CGGGCATGTT TCTGACTTTG TTTATTAAACA  
 12121 CTGTTTGAA CATCACCATA GCAAGCAGGG TACTGGAGCA GAGACTCACT GACTCCGCCT  
 12181 GTGCGGCCCTT CATCGGGCAGAC GACAACATCG TTCAACGGAGT GATCTCCGAC AAGCTGATGG  
 12241 CGGAGAGGTG CGCGTCGTGG GTCAACATGG AGGTGAAGAT CATTGACGCT GTCATGGCG  
 12301 AAAAACCCCC ATATTTTTGT GGGGGATTCA TAGTTTTGA CAGCGTCACA CAGACCGCCT  
 12361 GCCGTGTTTC AGACCCACTT AAGCGCCTGT TCAAGTTGGG TAAGCCGCTA ACAGCTGAAG  
 12421 ACAAGCAGGA CGAAGACAGG CGACGAGCAC TGAGTGAACGA GGTT

FIGURE 29E

68/240

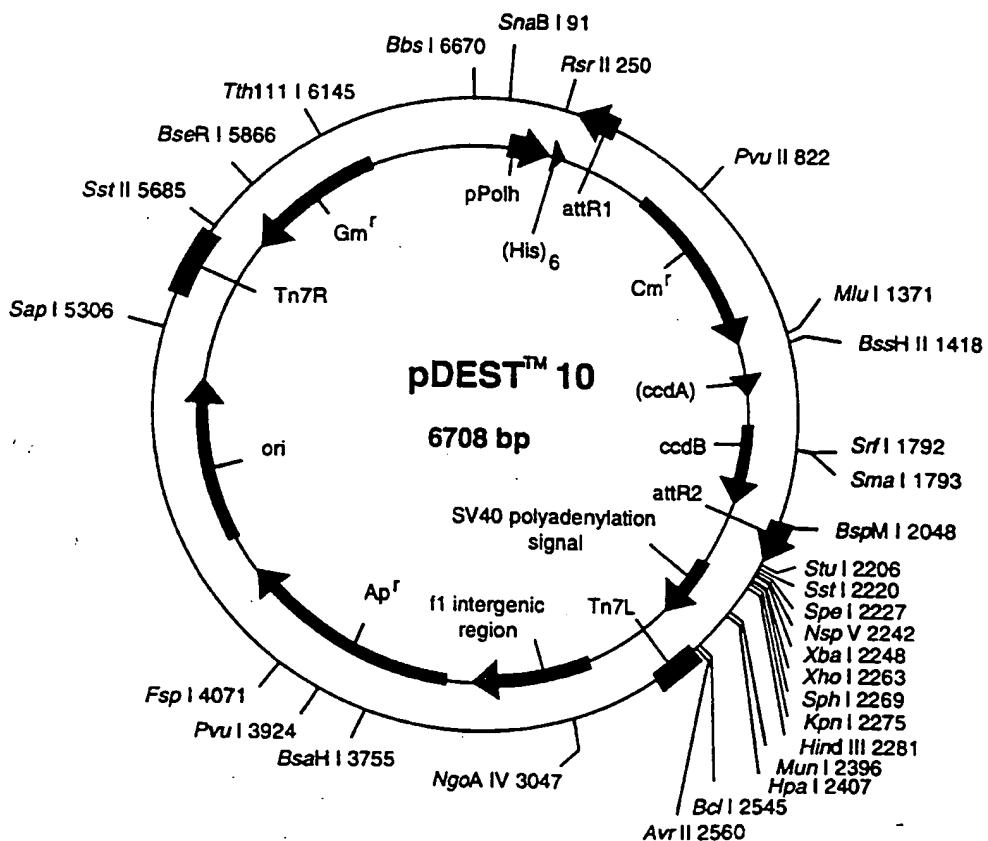
**Figure 30A:** pDEST10 Polyhedron Promoter with N-His6,  
Baculovirus Transfer Plasmid

154 *mRNA from polyhedrin promoter*  
 aaa taa gta ttt tac tgt ttt cgt aac agt ttt gta ata aaa aaa cct ata  
 ttt att cat aaa atg aca aaa gca ttg tca aaa cat tat ttt ttt gga tat

205 aat att ccg gat tat tca tac cgt ccc acc atc ggg cgc gga tct cgg tcc  
 tta taa ggc cta ata agt atg gca ggg tgg tag ccc gcg cct aga gcc agg

256 Met Ser Tyr Tyr His His His His His Asp Tyr Asp Ile Pro  
 gaa acc atg tcg tac tac cat cac cat cac cat cac gat tac gat atc cca  
 ctt tgg tac agc atg atg gta gtg gta gtg cta atg cta tag ggt

307 Thr Thr Glu Asn Lys Tyr Phe Gln<sup>+</sup> Glu Ile Thr Ser Lys Tyr Lys Lys  
 acg acc gaa aac ctg tat ttt cag ggc atc aca agt ttc/tsc aca gaa gct  
 tgc tgg ctt ttg gac ata aaa gtc ccg tag tgc tca aac atg ttc/tcc ggc  
 attR1 Int



69/240

## pDEST10 6708 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
23..152	Ppolh
461..337	attR1
711..1370	CmR
1490..1574	inactivated ccdA
1712..2017	ccdB
2058..2182	attR2
3394..4369	ampR
4510..5164	ori
5658..62	genR

1 CCCCGGATGA AGTGGTTCGC ATCCCTGGTT TTCTGGAGG CGAGCATCGT TTGTTGCC  
 61 AGGACTCTAG CTATAGTTCT AGTGGTTGC TACGTATACT CCGGAATATT AATAGATCAT  
 121 GGAGATAATT AAAATGATAA CCATCTCGCA AATAAATAAG TATTTTACTG TTTTCGTAAAC  
 181 AGTTTTGTAA TAAAAAAAACC TATAAATATT CCGGATTATT CATACCGTCC CACCATCGGG  
 241 CGCGGATCTC GGTCCGAAAC CATGTGTCAC TACCATCAC ACCACCATCA CGATTACGAT  
 301 ATCCCAACGA CCGAAAACCT GTATTTTCAG GGCATCACAA GTTTGTACAA AAAAGCTGAA  
 361 CGAGAAACGT AAAATGATAT AAATATCAAT ATATTAATT AGATTTTGCA TAAAAAACAG  
 421 ACTACATAAT ACTGTAAAAC ACAACATATC CAGTCACTAT GGCGGCCGCT AAGTTGCCAG  
 481 CATCACCCGA CGCACTTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC TTCGAGAAT  
 541 AAATAAATCC TGGTGTCCCT GTTGATACCG GGAAGCCCTG GGCCAACCTT TGGCGAAAAT  
 601 GAGACGTTGA TCGGCACGTA AGAGGTTCCA ACTTCACCA TAATGAAATA AGATCACTAC  
 661 CGGGCGTATT TTTTGAGTTA TCGAGATT T CAGGAGCTAA GGAAGCTAAA ATGGAGAAAA  
 721 AAATCACTGG ATATACCACC GTTGATATAT CCCAATGGCA TCGTAAAGAA CATTGGAGG  
 781 CATTTCACTGC AGTTGCTCAA TGTACCTATA ACCAGACCGT TCAGCTGGAT ATTACGGCCT  
 841 TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTATCC GGCCTTATT CACATTCTTG  
 901 CCCGCCTGAT GAATGCTCAT CGGAATTCC GTATGGCAAT GAAAGACGGT GAGCTGGTGA  
 961 TATGGGATAG TGGTACCCCT TGTTACACCG TTTTCCATGA GCAAACGTGAA ACGTTTTCAT  
 1021 CGCTCTGGAG TGAATACCAC GACGATTTC GGCAGTTCT ACACATATAT TCGCAAGATG  
 1081 TGGCGTGTAA CGGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTTATTGAG AATATGTTTT  
 1141 TCGTCTCAGC CAATCCCTGG GTGAGTTCA CCAGTTTGA TTTAACGTG GCCAATATGG  
 1201 ACAACTTCTT CGCCCCCGTT TTCACCATGG GCAAATATTA TACGCAAGGC GACAAGGTGC  
 1261 TGATGCCGCT GGGGATTCAAG GTTCATCATG CCGCTCTGTGA TGGCTTCAT GTCGGCAGAA  
 1321 TGCTTAATGA ATTACAACAG TACTGCGATG AGTGGCAGGG CGGGGCGTAA ACGCGTGGAT  
 1381 CGGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTGCG CGCTGATT T TCGGTATAA  
 1441 GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAG AGGTGTGCTA TGAAGCAGCG  
 1501 TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCATATA TGATGCAAT  
 1561 ATCTCCGGTC TGGTAAGCAC AACCATGCG AATGAAGGCC GTCGTCTGCG TGCCGAACGC  
 1621 TGGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTGCGCC GTTTTATTGAA AATGAACGGC  
 1681 TCTTTGCTG ACCAGAACAG GGACTGGTGA AATGCAGTTT AAGGTTTACA CCTATAAAAG  
 1741 AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA CGCCCCGGCG  
 1801 ACGGATGGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAAGTCT CCCGTGAACCT  
 1861 TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGCCAG  
 1921 TGTGCCGGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG AAAATGACAT  
 1981 CAAAAACGCC ATTAACCTGA TGTTCTGGGG AATATAATG TCAGGCTCCC TTATACACAG  
 2041 CCAGCTGCA GGTGACCAT AGTGACTGGA TATGTTGTGT TTTACAGTAT TATGTAGTCT  
 2101 GTTTTTTATG CAAAATCTAA TTTAATATAT TGATATTAT ATCATTTCAT GTTTCTCGTT  
 2161 CAGCTTCTT GTACAAAGTG GTGATGCCAT GGATCCGAA TTCAAAGGCC TACGTCGACG  
 2221 AGCTCAACTA GTGCGGCCGC TTTCGAATCT AGAGCCTGCA GTCTCGAGGC ATGCGGTACC  
 2281 AAGCTTGTG AGAAGTACTA GAGGATCATA ATCAGCCATA CCACATTG AGAGGTTTA  
 2341 CTTGCTTTAA AAAACCTCCC ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT  
 2401 GTTGTGTTA ACTTGTATT TGCAGCTTAT AATGGTTACA AATAAAGCAA TAGCATCACA  
 2461 AATTTCACAA ATAAAGCATT TTTTCACTG CATTCTAGTT GTGGTTTGTC CAAACTCATC  
 2521 AATGTATCTT ATCATGTCTG GATCTGATCA CTGCTTGAGC CTAGGAGATC CGAACCCAGAT  
 2581 AAGTGAATC TAGTTCCAAA CTATTTGTC ATTGTTAATT TTGCTATTAG CTTACGACGC-

Figure 308

70/240

2641 TACACCCAGT TCCCATCTAT TTTGTCACTC TTCCCTAAAT AATCCTTAAA AACTCCATT  
 2701 CCACCCCTCC CAGTTCCCAA CTATTTGTC CGCCCACAGC GGGGCATTT TCTTCCCTGTT  
 2761 ATGTTTTAA TCAAACATCC TGCCAACCTC ATGTGACAAA CCGTCATCTT CGGCTACTTT  
 2821 TTCTCTGTCA CAGAATGAAA ATTTTTCTGT CATCTCTTCG TTATTAATGT TTGTAATTGA  
 2881 CTGAATATCA ACGCTTATTT GCAGCCTGAA TGGCGAATGG GACGCGCCCT GTAGCGGCGC  
 2941 ATTAAGCGCG GCGGGTGTGG TGGTTACCGC CAGCGTGACC GCTACACTTG CCAGGCCCT  
 3001 AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTCCTCGCC ACGTTCGCCG GCTTCCCCG  
 3061 TCAAGCTCTA AATCGGGGGC TCCCTTAGG GTTCCGATTT AGTGTCTTAC GGCACCTCGA  
 3121 CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT  
 3181 TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTITAATAGT GGACTCTTGT TCCAAACTGG  
 3241 AACAAACACTC AACCCCTATCT CGGTCTATTCT TTGATTTA TAAGGGATT TGCCGATTTC  
 3301 GGCCTATTGG TTTAAAAAAATG AGCTGATTIA ACAAAAAATTT AACGCGAATT TTAACAAAAT  
 3361 ATTAACGTTT ACAATTTCAG GTGGCACTT TCAGGGAAAT GTGCGCGGAA CCCCTATTG  
 3421 TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT  
 3481 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCGCGT TCGCCCTTAT  
 3541 TCCCTTTTTG GCGGCATTTT GCCTTCCTGT TTGCTCAC CCAGAAACGC TGGTGAAGT  
 3601 AAAAGATGCT GAAGATCAGT TGGGTGACAG AGTGGGTTAC ATCGAACTGG ATCTCAACAG  
 3661 CGGTAAGATC CTTGAGAGTT TCGCCCCGA AGAACGTTT CCAATGATGA GCACTTTAA  
 3721 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTG  
 3781 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT  
 3841 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGTGCC ATAACCATGA GTGATAACAC  
 3901 TCGGGCAAC TTACTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTGCA  
 3961 CAACATGGGG GATCATGTA CTCGCCTTGA TCGTTGGAA CGGGAGCTGA ATGAAGCCAT  
 4021 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACTG TGCGCAAAC  
 4081 ATTAACTGGC GAACTACTTA CTCTAGCTT CCGGAAACAA TTAATAGACT GGATGGAGGC  
 4141 GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA  
 4201 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG  
 4261 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG  
 4321 AAATAGACAG ATCGCTGAGA TAGGTGCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA  
 4381 AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTAA AAAGGATCTA  
 4441 GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA  
 4501 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG  
 4561 CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGGTGGTTT GTTTGCAGGA  
 4621 TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACAAA  
 4681 TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACGCC  
 4741 TACATACCTC GCTCTGCTAA TCCCTGTTAC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG  
 4801 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC  
 4861 GGGGGTTCG TGACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC TGAGATACCT  
 4921 ACAGCGTGA CATTGAGAAA GCGCCACCGT TCCCGAAGGG AGAAAGGCAG ACAGGTATCC  
 4981 GGTAAAGCGGC AGGGTCGGA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCG  
 5041 GTATCTTAT AGTCCTGTCG GTTITCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG  
 5101 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCCTTTT TACGGTTCC  
 5161 GGCCTTTGCG TGGCTTTTG CTCACATGTT CTTCCTGCG TTATCCCTG ATTCTGTGGA  
 5221 TAACCGTATT ACCGCTTTG AGTGGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG  
 5281 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA  
 5341 TCTGTGCGGT ATTTCACACC GCAGACCAGC CGCGTAACCT GCGAAAATCG GTTACGGTTG  
 5401 AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA CAATAAAAGTC TTAAACTGAA  
 5461 CAAAATAGAT CTAAAATATG ACAATAAAAGT CTTAAACTAG ACAGAATAGT TGTAAACTGA  
 5521 AATCAGTCCA GTTATGCTGT GAAAAAAGCAT ACTGGACTTT TGTTATGGCT AAAGCAAAC  
 5581 CTTCATTTTC TGAAGTGCCTA ATTGCCCTGC GTATTAAGA GGGCGCTGGC CAAGGGCATG  
 5641 GTAAAGACTA TATTCGCGGC GTTGTACAA TTTACCGAAC AACTCCCGGG CCGGGAGGCC  
 5701 GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCATCAC  
 5761 TTCTTCCCGT ATGCCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC  
 5821 TTGCACGTAG ATCACATAAG CACCAAGCGC GTTGGCTCA TGCTTGAGGA GATTGATGAG  
 5881 CGCGGTGGCA ATGCCCTGCC TCCGGTGTCTC GCCGGAGACT GCGAGATCAT AGATATAGAT  
 5941 CTCACTACGC GGCTGCTCAA ACCTGGGAG AACGTAAGCC GCGAGAGCGC CAACAAACCGC  
 6001 TTCTTGGTCG AAGGCAGCAA GCGCGATGAA TGTCTTACTA CCGGAGCAAGT TCCCGAGGTA  
 6061 ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC GACCGAAAAG-

FIGURE 30C

71/240

6121 ATCAAGAGCA GCCCGCATGG AITTGACTTG GTCAGGGCCG AGCCTACATG TGCGAATGAT  
6181 GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCC TAACATCGTT  
6241 GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAACA TCGACCCACG GCGTAACGCG  
6301 CTTGCTGCTT GGATGCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA  
6361 AAACGCCAC TGCGCCGTTA CCACCGCTGC GTTCCGTCAA GGTTCTGGAC CAGTTGCGTG  
6421 AGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAACAA GGCTTATGTC AACTGGGTTC  
6481 GTGCCTTCAT CCGTTTCCAC GGTGTGCCAC ACCGGCAAC CTTGGGCAGC AGCGAAGTCG  
6541 AGGCATTCT GTCCCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG CATCGTCAGG  
6601 CATTGGCGGC CTTGCTGTTT TTCTACGGCA AGGTGCTGTG CACGGATCTG CCCTGGCTTC  
6661 AGGAGATCGG AAGACCTCGG CGTCGCCGGT GGTGCTGA

FIGURE 30D

72/240

Figure 31A: pDEST 11

## Tet-regulated eukaryotic expression

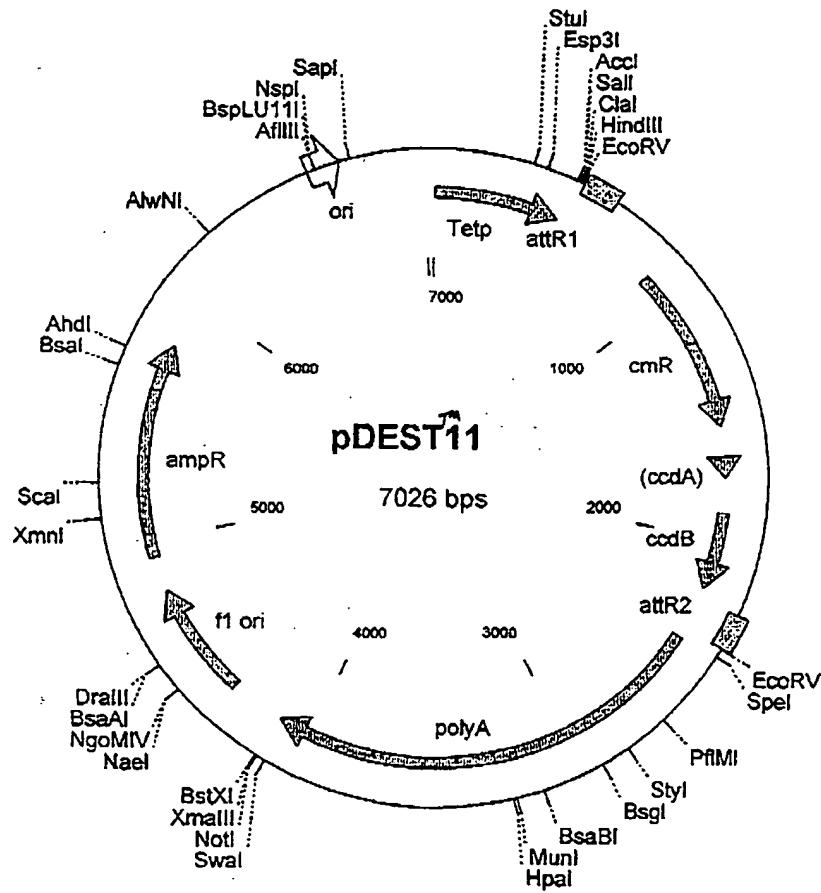
mRNA from CMV promoter (controlled by tetracycline)

358 tag tga acc gtc aga tcg cct gga gac gcc atc cac get gtt ttg acc tcc  
atc act tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg

409 ata gaa gac acc ggg acc gat cca gcc tcc gcg gcc ccg aat tcg agc tcg  
tat ctt ctg tgg ccc tgg cta ggt cgg agg cgc cgg ggc tta agc tcg agc

460 gta ccc ggg gat cct cta gag tcg agg tcg acg gta tcg ataa<sup>Sph</sup> tcg ttg aca  
cat ggg ccc cta gga gat ctc agc tcc agc ggc cat agc tat tcg aac tat  
Int attR1 Csp<sup>Hind</sup> 3 EcoRV

511 tca [aca agt ttg gag aat aat gct gaa cga gaa acg taa dat gat abt aat  
agt tgg tca aac atg ttt] ttg cgg ctt gct cta tgc att tta cta dat tta



73/240

## pDEST11 7026 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
4..479	Tet <sub>p</sub> ((Tet operator)7 and min hCMV promoter)
638..514	attR1
888..1547	CmR
1667..1751	inactivated ccdA
1889..2194	ccdB
2235..2359	attR2
2402..4132	polyA
4347..4803	f1 ori
4940..5797	ampR

1 CGAGTTTACC ACTCCCTATC AGTGATAGAG AAAAGTGAAG GTCGAGTTTA CCACTCCCTA  
 61 TCAGTGATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCGATAG AGAGAAAAGT  
 121 GAAAGTCGAG TTTACCACTC CCTATCGATG ATAGAGAAAA GTGAAAGTCG AGTTTACCA  
 181 TCCCTATCAG TGATAGAGAA AAGTGAAGT CGAGTTTACC ACTCCCTATC AGTGATAGAG  
 241 AAAAGTGAAGA GTCGAGTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGCT  
 301 CGGTACCCGG GTCGAGTAGG CGTGTACGGT GGGAGGCCCTA TATAAGCAGA GCTCGTTAG  
 361 TGAACCGTCA GATCGCCTGG AGACGCCATC CACGCTGTT TGACCTCCAT AGAACGACACC  
 421 GGGACCGATC CAGCCTCCGC GGCCCCGAAT TCGAGCTCGG TACCCGGGGA TCCTCTAGAG  
 481 TCGAGGTGCA CGGTATCGAT AAGCTTGATA TCAACAAGTT TGTACAAAAAA AGCTGAACGA  
 541 GAAACGTAAA ATGATATAAA TATCAATATA TTAAATTAGA TTTTGCATAA AAAACAGACT  
 601 ACATAATACT GTAAAACACA ACATATCCAG TCACTATGGC GGCCGCTAAG TTGGCAGCAT  
 661 CACCCGACGC ACTTTGCGCC GAATAAACAC CTGTGACGGG AGATCACTTC GCAGAATAAA  
 721 TAAATCCTGG TGTCCTGTT GATACCGGG AGCCCTGGGC CAACCTTTGG CGAAAATGAG  
 781 ACGTTGATCG GCACGTAAGA GGTTCCAAGT TTCACCCATAA TGAAATAAGA TCACTACCGG  
 841 GCGTATTTTG TGAGTTATCG AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA  
 901 TCACTGGATA TACCACCGTT GATATATCCC AATGGCATCG TAAAGAACAT TTGAGGCAT  
 961 TTCAGTCAGT TGCTCAATGT ACCTATAACC AGACCGTTCA GCTGGATATT ACGGCCTTTT  
 1021 TAAAGACCGT AAAGAAAAAT AAGCACAAGT TTTATCCGGC CTTTATTAC ATTCTTGCC  
 1081 GCCTGATGAA TGCTCATCCG GAATTCCGTA TGGCAATGAA AGACGGTGAG CTGGTGTAT  
 1141 GGGATAGTGT TOACCCCTGT TACACCGTT TCCATGAGCA AACTGAAACG TTTTCATCGC  
 1201 TCTGGAGTGA ATACCACGAC GATTTCGGC AGTTTCTACA CATATATTG CAAAGATGTGG  
 1261 CGTGTACCGG TGAAAACCTG GCCTATTTCCTA CTAAAGGGTT TATTGAGAAT ATGTTTTTCG  
 1321 TCTCAGCCAA TCCCTGGGTG AGTTTCACCA GTTTGATTT AAACGTGGCC AATATGGACA  
 1381 ACTTCTTCGC CCCCGTTTTC ACCATGGGCA AATATTATAC GCAAGGGCAC AAGGTGCTGA  
 1441 TGCCGCTGGC GATTCAAGGTT CATCATGCCG TCTGTGATGG CTTCCATGTC GGCAGAAATGC  
 1501 TTAATGAATT ACAACAGTAC TGCGATGAGT GGCAGGGCGG GGCTAAAGA TCTGGATCCG  
 1561 GCTTACTAAA AGCCAGATAA CAGTATGCGT ATTTGCGCGC TGATTTTG C GGTATAAGAA  
 1621 TATATACTGA TATGTATACC CGAAGTATGT CAAAAAGAGG TGTGCTATGA AGCAGCGTAT  
 1681 TACAGTGACA GTTGACAGCG ACAGCTATCA GTTGCTCAAG GCATATATGA TGTCAATATC  
 1741 TCCGGTCTGG TAAGCACAAC CATGCAAGAT GAAGCCCGTC GTCTGCGTGC CGAACGCTGG  
 1801 AAAGCGGAAA ATCAGGAAGG GATGGCTGAG GTGGCCGGT TTATTGAAAT GAACGGCTCT  
 1861 TTGCTGACG AGAACAGGGG CTGGTGAAT GCAGTTAAG GTTACACCT ATAAAAGAGA  
 1921 GAGCCGTTAT CGCTGTTG TGGATGTACA GAGTGATATT ATTGACACGC CGGGGGACG  
 1981 GATGGTGATC CCCCTGGCA GTGCACGTCT GCTGTCAGAT AAAGTCTCCC GTGAACTTTA  
 2041 CCCGGTGGTG CATATCGGG ATGAAAGCTG GCGCATGATG ACCACCGATA TGGCCAGTGT  
 2101 GCCGGTCTCC GTTATCGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA  
 2161 AAACGCCATT AACCTGATGT TCTGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA  
 2221 GTCTGCAGGT CGACCATAGT GACTGGATAT GTTGTTTACAGTATTAT GTAGTCTGTT  
 2281 TTTTATGCAA ATCTAATT AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCA  
 2341 CTTTCTTGTG CAAAGTGGT GATATCGGAAT TCCTGCAGCC CGGGGGATCC ACTAGTTCTA  
 2401 GAGCACTGCG ATGAGTGGCA GGGCGGGCG TAATTTTTT AAGGCAGTTA TTGGTGCCCT  
 2461 TAAACGCCTG GTGCTACGCC TGAATAAGTG ATAATAAGCG GATGAATGGC AGAAATTGCG  
 2521 CGGATCTTG TGAAGGAACC TTACTTGTG GGTGTGACAT AATTGGACAA ACTACCTACA-

FIGURE 31B

74/740

2581 GAGATTTAAA GCTCTAAGGT AAATATAAAA TTTTTAAGTG TATAATGTGT TAAACTACTG  
 2641 ATTCTAATTG TTTGTGTATT TTAGATTCCA ACCTATGGAA CTGATGAATG GGAGCAGTGG  
 2701 TGGAATGCCT TTAATGAGGA AAACCTGTT TGCTCAGAAG AAATGCCATC TAGTGATGAT  
 2761 GAGGCTACTG CTGACTCTCA ACATTCTACT CCTCCAAAAA AGAAGAGAAA GGTAGAAGAC  
 2821 CCCAAGGACT TTCCCTCAGA ATTGCTAAGT TTTTGAGTC ATGCTGTGTT TAGTAATAGA  
 2881 ACTCTTGCTT GCTTGTCTAT TTACACCACA AAGGAAAAAG CTGCACTGCT ATACAAGAAA  
 2941 ATTATGGAAA AATATTCTGT AACCTTTATA AGTAGGCATA ACAGTTATAA TCATAACATA  
 3001 CTGTTTTTC TTACTCCACA CAGGCATAGA GTGCTGCTA TTAATAACTA TGCTCAAAAA  
 3061 TTGTTGACCT TTAGCTTTT AATTTGAAA GGGGTTAATA AGGAATATTT GATGTATAGT  
 3121 GCCTTGACTA GAGATCATAA TCAGCCATAC CACATTGTA GAGGTTTAC TTGCTTTAAA  
 3181 AAACCTCCCA CACCTCCCCC TGAAACCTGAA ACATAAAATG AATGCAATTG TTGTTGTTAA  
 3241 CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAGCAAT AGCATCACAA ATTCACAAA  
 3301 TAAAGCATT TTTTCACTGC ATTCTAGTTG TGGTTGTC AAACATCATCA ATGTATCTTA  
 3361 TCATGTCCTGG ATCCCCAGGA AGCTCCTCTG TGTCCTCATA AACCTAACCC TCCTCTACTT  
 3421 GAGAGGACAT TCCAATCATA GGCTGCCAT CCACCCCTCTG TGTCCTCCTG TTAATTAGGT  
 3481 CACTTAACAA AAAGGAAATT GGGTAGGGT TTTTCACAGA CCGCTTTCTA AGGGTAATT  
 3541 TAAAATATCT GGGAAAGTCCC TTCCACTGCT GTGTTCCAGA AGTGTGGTA AACAGCCAC  
 3601 AAATGTCAAC AGCAGAAACA TACAAGCTGT CAGCTTGCA CAAGGGCCCA ACACCCCTGCT  
 3661 CATCAAGAAG CACTGTGGTT GCTGTGTTAG TAATGTGCAA AACAGGAGGC ACATTTTCCC  
 3721 CACCTGTGA GTGTTCAAAA TATCTAGTGT TTTCATTTT ACTTGGATCA GGAACCCAGC  
 3781 ACTCCACTGG ATAAGCATTAA TCCTTATCCA AAACAGCCCT GTGGTCAGTG TTCATCTGCT  
 3841 GACTGTCAAC TGAGTCATT TTTGGGGTTA CAGTTTGAGC AGGATATTG GTCTGTAGT  
 3901 TTGCTAACAC ACCCTGCAGC TCCAAAGGTT CCCACCAAC AGCAAAAAAA TGAAAATTG  
 3961 ACCCTTGAAT GGGTTTCCA GCACCATTTT CATGAGTTTT TTGTTGTCCT GAATGCAAGT  
 4021 TTAACATAGC AGTTACCCCA ATAACCTCAG TTTAACAGT AACAGCTTCC CACATCAAAA  
 4081 TATTTCCACA GTGTTAAGTCC TCATTTAAAT TAGGCAAAGG AATTGCTCTA GAGCGGCCGC  
 4141 CACCGCGGTG GAGCTCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG  
 4201 TCGTTTTACA ACGTGCGTGC TGGGAAAACC CTGGCGTTAC CCAACTTAAAT CGCCTTGCAG  
 4261 CACATCCCCC TTTCGCCAGC TGGCGTAATA GCGAAGAGGC CCGCACCGAT CGCCCTTCCC  
 4321 AACAGTTGCG CAGCCTGAAT GGCAGATGGG ACGCGCCCTG TAGCGCGCA TTAAGCGCGG  
 4381 CGGGTGTGGT GTGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC  
 4441 CTTTCGCTTT CTTCCCTTCC TTTCTGCCA CGTTGCCGG CTTTCCCCGT CAAGCTCTAA  
 4501 ATCGGGGGCT CCCTTTAGGG TTCCGATTAA GTGTTTACG GCACCTCGAC CCCAAAAAAC  
 4561 TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCTT  
 4621 TGACCTTGGGA GTCCACGTTT TTAAATAGTG GACTCTTGTG CCAAACATGGA ACAACACTCA  
 4681 ACCCTATCTC GGTCTATTCT TTTGATTITAT AAGGGATTTT GCCGATTTCG GCCTATTGGT  
 4741 TAAAAATGAA GCTGATTAA CAAAAATTAA ACGCGAATT TAAACAAAATA TTAACGCTTA  
 4801 CAATTAGGT GGCACTTTTC GGGGAAATGT GCGCGGAACC CCTATTGTT TATTTTCTA  
 4861 AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAAATGC TTCAATAATA  
 4921 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCTGTC GCCCTTATTC CCTTTTTGTC  
 4981 GGCATTTGC CTTCCGTGTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA  
 5041 AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACCTGGAT CTCAACAGCG GTAAGATCCT  
 5101 TGAGAGTTT CGCCCCGAAG AACGTTTTC AATGATGAGC ACTTTAAAG TTCTGCTATG  
 5161 TGGCCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA  
 5221 TTCTCAGAAT GACTTGGTT AGTACTCACC AGTCACAGAA AAGCATCTTA CGGATGGCAT  
 5281 GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCAGTGGT GATAACACTG CGGCCAACTT  
 5341 ACTTCGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGACACA ACATGGGGGA  
 5401 TCATGTAACCT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA  
 5461 GCGTGACACC ACGATGCCCTG TAGCAATGGC AACAACGTT CGCAAACATAT TAACTGGCGA  
 5521 ACTACTTACT CTAGCTTCCC GGCAACAAATT AATAGACTGG ATGGAGGGCGG ATAAAGTTGC  
 5581 AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC  
 5641 CGGTGAGCGT GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG  
 5701 TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT  
 5761 CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGTTAAGTG TCAGACCAAG TTACTCATA  
 5821 TATACTTTAG ATTGATTTAA AACTTCATT TTAATTAAA AGGATCTAGG TGAAGATCCT  
 5881 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCCTTCCACT GAGCGTCAGA  
 5941 CCCCCTAGAA AAGATCAAAG GATCTTCTTG AGATCCTTTT TTCTGCCGG TAATCTGCTG  
 6001 CTTGCAAACA AAAAAACCAC CGCTACCAGC GGTGGTTTGT TTGCCGGATC AAGAGCTACC-

FIGURE 3/C

75/240

6061 AACTTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT  
6121 AGTGTAAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATAACCTCGC  
6181 TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT  
6241 GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTG GGCTGAACGG GGGGTTCTG  
6301 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAAGTG AGATACTAC AGCGTGAGCT  
6361 ATGAGAAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCAGC AGGTATCCGG TAAGCGGCAG  
6421 GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGAA AACGCCTGGT ATCTTTATAG  
6481 TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG  
6541 GCGGAGCCTA TGAAAAAACG CCAGCAACGC GGCTTTTTA CGGTTCCCTGG CCTTTTGCTG  
6601 GCCTTTGCTCACATGTTCT TTCTCGCGTT ATCCCCGTAT TCTGTGGATA ACCGTATTAC  
6661 CGCCTTGAG TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT  
6721 GAGCGAGGAA GCGGAAGAGC GCCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT  
6781 TCATTAATGCA AGCTGGCACG ACAGGTTCC CGACTGGAAA GCGGGCAGTG AGCGCAACGC  
6841 AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCCGGC  
6901 TCGTATGTTG TGTTGGAAATTG TGAGCGGATA ACAATTTCAC ACAGGAAACA GCTATGACCA  
6961 TGATTACGCC AAGCGCGCAA TTAACCCCTCA CTAAAGGGAA CAAAAGCTGG GTACCGGGCC  
7021 CCCCT

FIGURE 31D

76/260

**Figure 32A: pDEST12.2 CMV Promoter for Eukaryotic Expression, SV40 Promoter/ori for G418 Resistance**

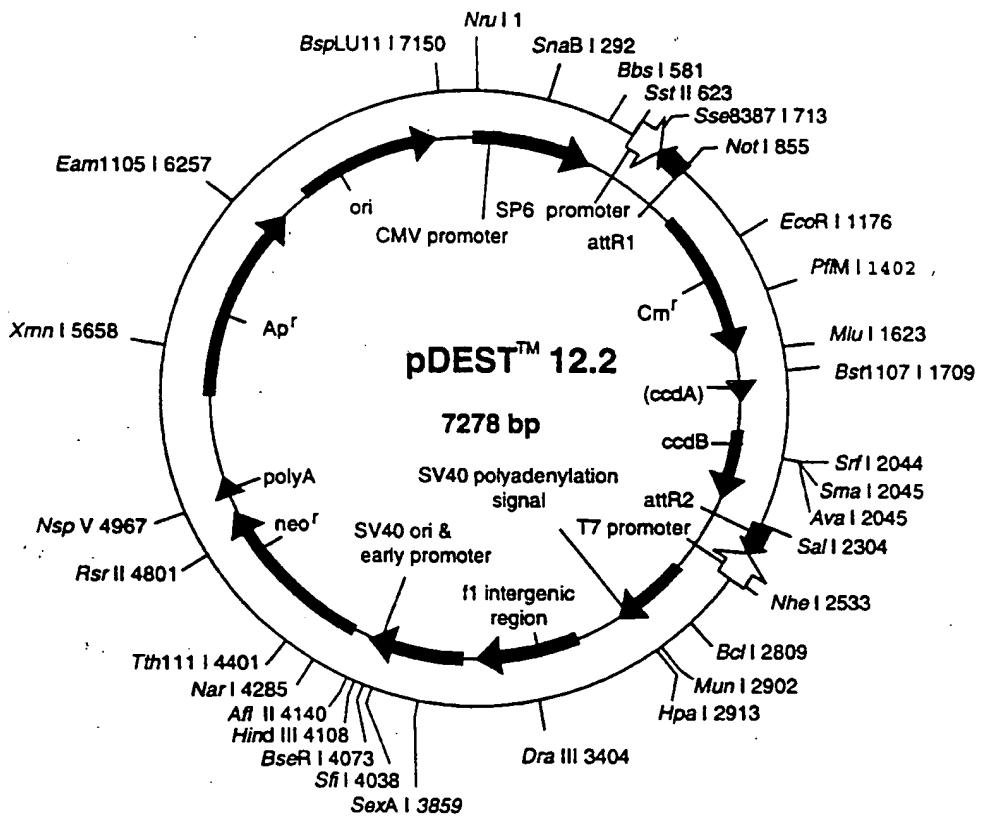
307 acc <sup>→ mRNA from CMV promoter</sup> gtc aga tcg cct gga gac gcc atc cac gct gtt ttg acc tcc ata gaa  
 tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt

358 gac acc ggg acc gat cca gcc tcc gga ctc tag cct agg ccc cgg agc gga  
 ctg tgg ccc tgg cta ggt cgg agg cct gag atc gga tcc ggc gcc tcc cct

409 taa caa ttt cac aca gga aac agc tat gac cat tag gcc ttt gca aaa agc  
 att gtt aaa gtt tgt cct ttg tcc ata ctg gta atc cgg aaa cgt ttt tcc

460 tat tta ggt gac act ata gaa ggt acg cct gca ggt <sup>Apa I</sup> <sup>EcoR I</sup> atc ggt ccc gaa ttc  
 ata aat cca ctg tga tat ctt cca tgc gga cgt cca tgg cca ggc ctt aag

511 cca tca <sup>Int</sup> <sup>attR1</sup> ~~aca agt ttg tao ada ada gct gaa cga gaa acg taa aat gat ata~~  
~~gtt agt tgt tca aac atg ttt tbt cga ctc gct ctt tgc att gta qta tat~~



77/240

## pDEST12.2 7278 bp (rotated to position 3900)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
86..136	ori
220..742	CMV promoter
1059..935	attR1
1168..1827	CmR
1947..2031	inactivated ccdA
2169..2474	ccdB
2515..2639	attR2
2824..3186	small t & polyA
3310..3378	lac
4363..5157	neo
5680..6540	ampR

1 GGGGGCGGAA GCCTATGGAA AAACGCCAGC AACGCCGCCT TTTTACGGTT CCTGGCCTTT  
 61 TGCTGCCCT TTGCTCACAT GTTCTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT  
 121 ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG  
 181 TCAGTGAGCG AGGAAGCGGA AGAGCTCGCG AATGCATGTC GTTACATAAC TTACGGTAAA  
 241 TGGCCCGCCT GGCTGACCAC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT  
 301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA  
 361 AACTGCCAC TTGGCAGTAC ATCAAGTGT A CATATGCCA AGTACGCCCT CTATTGACGT  
 421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAAGTAC ATGACCTTAT GGGACTTTCC  
 481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGTGAC GGTGTTGGCA  
 541 GTACATCAAT GGGCGTGGAT AGCGGTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT  
 601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA  
 661 CAACTCCGCC CCATTGACGC AAATGGCGG TAGGCGTGTAA CGGTGGGAGG TCTATATAAG  
 721 CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTGACCT  
 781 CCATAGAAGA CACCGGGACC GATCCAGCCT CGGCACTCTA GCCTAGGCCG CGGGACGGAT  
 841 ACAAAATTCA CACAGGAAAC AGCTATGACC ATTAGGCCTT TGCAAAAAGC TATTTAGGTG  
 901 ACACATATAGA AGGTACGCCT GCAGGTACCG GATCACAAGT TTGTACAAAA AAGCTGAACG  
 961 AGAAACGTAA AATGATATAA ATATCAATAT ATAAATTAG ATTTTGCATA AAAAACAGAC  
 1021 TACATAATAC TGAAAACAC AACATATCCA GTCACATATGG CGGCCGCATT AGGCACCCCA  
 1081 GGCTTACAC TTATGCTTC CGGCTCGTAT AATGTGTGGA TTTTGAGTTA GGATCCGTCG  
 1141 AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA TCACTGGATA TACCAACCGTT  
 1201 GATATATCCC AATGGCATCG TAAAGAACAT TTTGAGGCAT TTCAGTCAGT TGCTCAATGT  
 1261 ACCTATAACC AGACCGTCA GCTGGATATT ACGGCCTTTT TAAAGACCGT AAAGAAAAT  
 1321 AAGCACAAGT TTATCCGGC CTTTATTCCAC ATTCTTGCCTT GCCTGATGAA TGCTCATCCG  
 1381 GAATTCGTA TGGCAATGAA AGACGGTGAG CTGGTGATAT GGGATAGTGT TCACCCCTTGT  
 1441 TACACCGTT TCCATGAGCA AACTGAAACG TTTTCATCGC TCTGGAGTGA ATACCACGAC  
 1501 GATTTCCGGC AGTTTCTACA CATATATTCCG CAAGATGTGG CGTGTACCGG TGAAAACCTG  
 1561 GCCTATTCC CTAAGGGTT TATTGAGAAT ATGTTTTCG TCTCAGCCAA TCCCTGGGTG  
 1621 AGTTTCACCA GTTTGATTT AAACGTGGCC AATATGGACA ACTTCTTCGC CCCCCGTTTTC  
 1681 ACCATGGCA AATATTATAC GCAAGGCAC AAGGTGCTGA TGCCGCTGGC GATTCAAGGTT  
 1741 CATCATGCCG TCTGTGATGG CTTCCATGTC GGCAGAATGCA TTAATGAATT ACAACAGTAC  
 1801 TGCAGATGAGT GGCAGGGCGG GGCAGTAAACG CGTGGATCCG CTCTACTAAA AGCCAGATAA  
 1861 CAGTATGCGT ATTTGCGCGC TGATTTTGTG GGTATAAGAA TATATACTGA TATGTATACC  
 1921 CGAAGTATGT CAAAAGAGG TGTGCTATGA AGCAGCGTAT TACAGTGACA GTTGACAGCG  
 1981 ACAGCTATCA GTTGCTCAAG GCATATATGA TGTCAATATC TCCGGTCTGG TAAGCACAAC  
 2041 CATGCAGAAT GAAGCCCGTC GTCTGCGTGC CGAACGCTGG AAAGCGGAAA ATCAGGAAGG  
 2101 GATGGCTGAG GTGCCCGGGT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACAGGGA  
 2161 CTGGTAAAT GCAAGTTAACG GTTTACACCT ATAAAAGAGA GAGCCGTTAT CGTCTGTTTG  
 2221 TGGATGTACA GAGTGTATT ATTGACACGC CGGGCGACG GATGGTGTAC CCCCTGGCCA  
 2281 GTGCACGTCT GCTGTCAAGAT AAAGTCTCCC GTGAACTTTA CCCGGTGGTG CATATCGGGG  
 2341 ATGAAAGCTG GCGCATGATG ACCACCGATA TGGCCAGTGT GCCGGTCTCC GTTATCGGGG  
 2401 AAGAAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA AAACGCCATT AACCTGATGT-

FIGURE 32B

78/240

2461 TCTGGGAAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATACT  
 2521 GACTGGATAT GTTGTGTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATT  
 2581 AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCA CTTTCTTGTA CAAAGTGGTG  
 2641 ATCGCGTGC A TGCGACGTCA TAGCTCTTC CCTATAGTGA GTCGTATTAT AAGCTAGGCA  
 2701 CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAAGTGT AGCTTGGGAT CTTTGTGAAG  
 2761 GAACCTTACT TCTGTGGTGT GACATAATTG GACAAACTAC CTACAGAGAT TTAAAGCTCT  
 2821 AAGGTTAAATA TAAAATTTT AAGTGTATAA TGTGTTAAC TAGCTGCATA TGCTTGTGTC  
 2881 TTGAGAGTTT TGCTTACTGA GTATGATT TA TGAAATATT ATACACAGGA GCTAGTGATT  
 2941 CTAATTGTTT GTGTATTITA GATTCACTAGT CCCAAGGCTC ATTTCAAGGCC CCTCAGTCCT  
 3001 CACAGTCTGT TCATGATCAT AATCAGCCAT ACCACATTG TAGAGGTTTT ACTTGCTTTA  
 3061 AAAAACCTCC CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTTGTGTT  
 3121 AACTTGTATA TTGCACTGTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTCACA  
 3181 AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGTTTGT CCAAACACTCAT CAATGTATCT  
 3241 TATCATGTCT GGATCGATCC TGCAATTAG AATCGGCCAA CGCGCGGGGA GAGGCGGGTT  
 3301 GCGTATTGGC TGCGTAATA GCGAAGAGGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG  
 3361 CAGCCTGAAT GCGGAATGGG ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT  
 3421 GGTTACGCGC AGCGTGACCG CTACACTTGC CAGGCCCTA GCGCCCGCTC CTTTCGCTTT  
 3481 CTTCCCTTCC TTTCTGCCA CGTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT  
 3541 CCCTTTAGGG TTCCGATTTA GTGCTTACG GCACCTCGAC CCCAAAAAAAC TTGATTAGGG  
 3601 TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCT TGACGTTGGA  
 3661 GTCCACGTT TTTAATAGTG GACTCTGTT CCAAACACTGGA ACAACACTCA ACCCTATCTC  
 3721 GGTCTATTCT TTGATTAT AAGGGATT TTGATTTG GCCGATTTG GCCTATTGGT TAAAAAAATGA  
 3781 GCTGATTTAA CAAATATT TAACAAAATA TIAACGTTTA CAATTTCGCC  
 3841 TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTC ACACCGCATA CGCGGATCTG  
 3901 CGCAGCACCA TGCCCTGAAA TAACCTCTGA AAGAGGAAC TGGTTAGGTA CTTTCTGAGG  
 3961 CGGAAAGAAC CAGCTGTGGA ATGTGTGCA GTAGGGTGT GGAAAGTCCC CAGGCTCCCC  
 4021 AGCAGGCAGA AGTATGCAA GCATGCATCT CAATTAGTCA GCAACCAGGT GTGGAAAGTC  
 4081 CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT CTCAATTAGT CAGCAACCAT  
 4141 AGTCCCGCCC CTAACCTCCGC CCATCCCGCC CCTAACTCCG CCCAGTTCCG CCCATTCTCC  
 4201 GCCCCATGGC TGACTAATT TTTTATTGTA TGCAAGGCC GAGGCCGCCT CGGCCTCTGA  
 4261 GCTATTCCAG AAGTAGTGAG GAGGCTTTT TGGAGGCCA GGCTTTTGCA AAAAGCTTGA  
 4321 TTCTTCTGAC ACAACAGTCT CGAACTTAAG GCTAGAGCCA CCATGATTGA ACAAGATGGA  
 4381 TTGCACTGAG GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCGGCTATGA CTGGGCACAA  
 4441 CAGACAATCG GCTGCTCTGA TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT  
 4501 CTTTTGTCA AGACCGACCT GTCCGGTGC CTGAATGAAC TGCAGGACGA GGCAGCGCG  
 4561 CTATCGTGGC TGCCACGAC GGGCGTTCT TGCGCAGCTG TGCTCGACGT TGTCACTGAA  
 4621 GCGGGAAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC AGGATCTCCT GTCATCTCAC  
 4681 CTTGCTCCTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TGCGGCCGGCT GCATACGCTT  
 4741 GATCCGGCTA CCTGCCATT CGACCACAA GCGAACACATC GCATCGAGCG AGCACGTACT  
 4801 CGGATGGAAG CCGGTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA GGGGCTCGCG  
 4861 CCAGCCGAAC TGGTCGCCAG GCTCAAGGCC CGCATGCCCG ACGGCGAGGA TCTCGTGTG  
 4921 ACCCATGGCG ATGCCCTGCTT GCGAATATC ATGGGGAAA ATGGCCGCTT TTCTGGATT  
 4981 ATCGACTGTG GCCGGCTGGG TGTGGCGAC CGCTATCAGG ACATAGCGTT GGCTACCGT  
 5041 GATATTGCTG AAGAGCTTGG CGCGAATGG GCTGACCGCT TCCCTCGTGT TTACGGTATC  
 5101 GCGCCTCCG ATTGCGAGCG CATCGCCTTC TATGCCCTTC TTGACGAGTT CTTCTGAGCG  
 5161 GGACTCTGGG GTTCGAAATG ACCGACCAAG CGACGCCAA CCTGCCATCA CGATGGCCGC  
 5221 AATAAAATAT CTTTATTTC ATTACATCG TGTGTTGGTT TTTTGTGTGA ATCGATAGCG  
 5281 ATAAGGATCC GCGTATGGTG CACTCTCACT ACAATCTGCT CTGATGCCGC ATAGTTAAC  
 5341 CAGCCCGAC ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGCA  
 5401 TCCGCTTACA GACAAGCTGT GACCGCTCC GGGAGCTGCA TGTGTCAAGAG GTTTTACCG  
 5461 TCATCACCGA AACCGCGCGAG ACGAAAGGC CTCGTGATAC GCCTATTGTT ATAGGTTAAT  
 5521 GTCATGATAA TAATGGTTT TTAGACGTCA GGTGGCACTT TTGGGGAAA TGTGCAGCG  
 5581 ACCCTATTGTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA  
 5641 CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT  
 5701 GTCGCCCTTA TTCCCTTTTG TGCGGCATTT TGCCCTCCTG TTTTGTGTCA CCCAGAAAC  
 5761 CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG  
 5821 GATCTCAACA GCGGTAAAGAT CCTTGAGAGT TTGCGCCCG AAGAACGTTT TCCAATGATG  
 5881 AGCACTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG-

FIGURE 32C

79/240

5941 CAACTCGGTC GCCGCATAACA CTATTCTAG AATGACTTGG TTGAGTACTC ACCAGTCACA  
6001 GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCAGT  
6061 AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAAC  
6121 GCTTTTTGCA ACAACATGGG GGATCATGTA ACTCGCCCTG ATCGTTGGGA ACCGGAGCTG  
6181 AATGAAGCCA TACCAAACGA CGAGCGTAC ACCACGATGC CTGTAGCAAT GGCAACAAACG  
6241 TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC  
6301 TGGATGGAGG CGGATAAAAGT TGCAAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG  
6361 TTTATTGCTG ATAAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG  
6421 GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT  
6481 ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGGCCT CACTGATTAA GCATTGGTAA  
6541 CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATT  
6601 AAAAGGATCT AGGTGAAGAT CTTTTTGAT AATCTCATGA CAAAATCCC TTAACGTGAG  
6661 TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT  
6721 TTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT  
6781 TGTGCGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG  
6841 CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT  
6901 GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC  
6961 GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGGATAGT TACCGGATAA GGCGCAGCGG  
7021 TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA  
7081 CTGAGATACC TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG  
7141 GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGG GCTTCCAGGG  
7201 GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA  
7261 TTTTTGTGAT GCTCGTCA

FIGURE 32D

80/260

Figure 33A: pDEST13

Native protein in E. coli:  $\lambda$ PL  
promoter

*BglII*

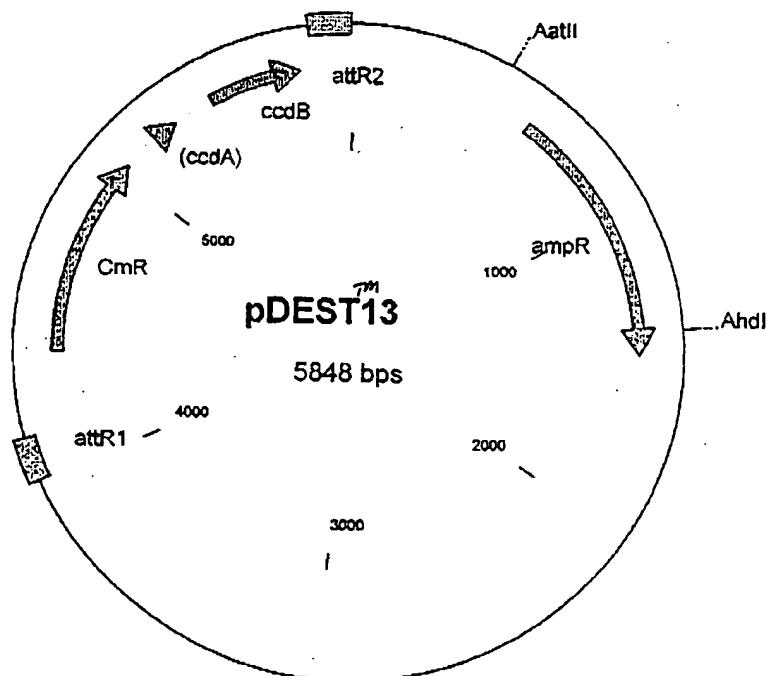
3721 tgggcaaacc aagacagcta aagatctctc acctacccaa caatgc~~cccc~~ ctgcaaaaaa  
acccgttgg ttctgtcgat ttcttagagag tggatggttt gttacgggg gacgtttttt

3781 taaaattcata taaaaaacat acagataacc atctgcggtg ataaattatc tctggcggtg  
attttaagtat attttttta tgcttattgg tagacccac tat~~taat~~ agaccgcac  
-35       $\lambda$ PL Promoter      -10      → mRNA

3841 ttgacataaa taccactggc ggtgatactg agcacatcg caggacgcac tgaccacat  
aactgtattt atggtgaccg ccactatgac tcgttagtc gtcctgcgtg actgggtgta  
EcoNI

3901 gaaggtgacg ctcttaaaaa ttaageccctg aagaaggca gcattcaaag cagaaggctt  
cttccactgc gagaattttt aattcgggac ttcttccgt cgtaagttc gtcttccgaa

3961 tgggtgtgt gatacggaaac gaagcattgg gatcatcaca agtttgtaca aaaaagctga  
accccacaca ctatgcttg cttcgttaacc ctatgtgt tcaaacatgt ttttcgact,



81/240

## pDEST13 5848 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
599..1458	ampR
4123..3998	attR1
4372..5031	CmR
5151..5235	inactivated ccdA
5373..5678	ccdB
5719..5843	attR2

1 TTCACTGGCC GTCGTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA  
 61 TCGCCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA  
 121 TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC GGTATTTCT  
 181 CCTTACGCAT CTGTGCGGTA TTTCACACCG CATATGGTGC ACTCTCAGTA CAATCTGCTC  
 241 TGATGCCGCA TAGTTAACGC AGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG  
 301 GGCTTGTCTG CTCCCAGCAT CCGCTTACAG ACAAGCTGTG ACCGTCCTCG GGAGCTGCAT  
 361 GTGTCAGAGG TTTTACCGT CATCACCGAA ACGCGCGAGA CGAAAGGGCC TCGTGATACG  
 421 CCTATTTTA TAGGTTAATG TCATGATAAT AATGGTTCT TAGACGTCAG GTGGCACTTT  
 481 TCGGGGAAAT GTGCGCGGAA CCCCTATTG TTATTTTTC TAAATACATT CAAATATGTA  
 541 TCCGCTCATG AGACAATAAC CCTGATAAAAT GCTTCATAAA TATTGAAAAA GGAAGAGTAT  
 601 GAGTATTCAA CATTTCCTGT TCGCCCTTAT TCCCTTTTTT GCGGCATTTC GCCTCCCTGT  
 661 TTTGCTCAC CCGAAACCCG TGGTAAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG  
 721 AGTGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCGA  
 781 AGAACGTTT CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGGG TATTATCCCG  
 841 TATTGACGCC GGGCAAGAGC AACTCCGTG CGGCATACAC TATTCTCAGA ATGACTTGGT  
 901 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG  
 961 CAGTGTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG  
 1021 AGGACCGAAG GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCCTGA  
 1081 TCGTTGGAA CCGGAGCTGA ATGAAGCCAT ACCAACGAC GAGCGTGACA CCACGATGCC  
 1141 TGTAGCAATG GCAACAACGT TGCGCAAACCT ATTAACTGGC GAAACTACTTA CTCTAGCTTC  
 1201 CCGGCAACAA TTATAGACT GGATGGAGGC GGATAAAGTT GCAGGACAC TTCTGCGCTC  
 1261 GGCCCTTCCG GCTGGCTGGT TTATGCTGA TAAATCTGGA GCGGGTGAGC GTGGGTCTCG  
 1321 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC  
 1381 GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGTGCCTC  
 1441 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATT  
 1501 AAAACTTCAT TTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC  
 1561 CAAAATCCCT TAACGTAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA  
 1621 AGGATCTCT TGAGATCCCT TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC  
 1681 ACCGCTACCA GCGGTGGTTT GTTTGCCTGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT  
 1741 AACTGGCTTC AGCAGAGCGC AGATACAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG  
 1801 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCITTTACC  
 1861 AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACCATACTT  
 1921 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGG  
 1981 GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT  
 2041 TCCCGAAGGG AGAAAGCGG ACAGGTATCC GGTAAAGCGGC AGGGTGGAA CAGGAGAGCG  
 2101 CACGAGGGAG CTTCCAGGGG GAAACGCTG GTATCTTAT AGTCCTGTCG GGTITCGCCA  
 2161 CCTCTGACTT GAGCGTCGAT TTTTGTATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA  
 2221 CGCCAGCAAC GCGGCCTTTT TACGGTTCCCT GGCCTTTGCG TGGCCTTTTG CTCACATGTT  
 2281 CTTTCCCTGCG TTATCCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA  
 2341 TACCGCTCGC CGCAGCGAA CGACCGAGCG CAGCGAGTC GTGAGCGAGG AAGCGGAAGA  
 2401 GCGCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCAATTAAAT GCAGCTGGCA  
 2461 CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGAAC GCAATTAAAT TGAGTTAGCT  
 2521 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT  
 2581 TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGG  
 2641 CTGCGAGGTGA TGATTATCAG CCAGCAGAGA TTAAGGAAA CAGACAGGTT TATTGAGCGC  
 2701 TTATCTTCC CTTTATTTT GCTGCGTAA GTGCATAAA AACCATCTT CATAATTCAA-

FIGURE 33B

82/240

2761 TCCATTTACT ATGTTATGTT CTGAGGGGAG TGAAAATTCC CCTAATTCGA TGAAGATTCT  
 2821 TGCTCAATTG TTATCAGCTA TGCGCCGACC AGAACACCTT GCCGATCAGC CAAACGTCTC  
 2881 TTCAGGCCAC TGACTAGCGA TAACTTCCC CACAACGGAA CAACTCTCAT TGCACTGGAT  
 2941 CATTGGGTAC GTGGGTTTA GTGGTTGTA AAACACCTGA CCGCTATCCC TGATCAGTTT  
 3001 CTTGAAGGTA AACTCATCA CCCAAGTCT GGCTATGCAG AAATCACCTG GCTCAACAGC  
 3061 CTGCTCAGGG TCAACGAGA TTAACATTC GTCAGGAAAG CTTGGCTTGG AGCCTGTTGG  
 3121 TGCGGTCTAG GAATTAACCTT CAACCTCAAG CCAGAATGCA GAATCACTGG CTTTTTGGT  
 3181 TGTGCTTACCA CATCTCTCCG CATCACCTT GTAAAGGTT CTAAGCTTAG GTGAGAACAT  
 3241 CCCTGCCTGA ACATGAGAAA AAACAGGGTA CTCATACTCA CTTCTAAGTG ACGGCTGCAT  
 3301 ACTAACCGCT TCATACATCT CGTAGATTTCT TCTGGCGATT GAAGGGCTAA ATTCTTCAAC  
 3361 GCTAACTTTG AGAATTTTG CAAGCAATGC GGGGTATAA GCATTAATG CATTGATGCC  
 3421 ATTAATAAAA GCACCAACGC CTGACTGCC CATAACCCATC TTGCTGCGA CAGATTCTG  
 3481 GGATAAGCCA AGTTCACTTT TCTTTTTTC ATAAATTGCT TTAAGGCGAC GTGCGCTCTC  
 3541 AAGCTGCTCT TGTGTTAATG GTTTCTTTT TGTGCTCATA CGTTAAATCT ATCACCGCAA  
 3601 GGGATAAAATA TCTAACACCG TGCGTGTGTA CTATTTTAC TCTGGCGGTG ATAATGGTTG  
 3661 CATGACTAA GGAGGTTGTA TGGAAACACG CATAACCCCTG AAAGATTATG CAATGCGCTT  
 3721 TGGGCAAACC AAGACAGCTA AAGATCTCTC ACCTACCAAA CAATGCCCTC CTGCAAAAAA  
 3781 TAAATTCTATA TAAAAAAACAT ACAGATAACC ATCTGCGGTG ATAAATTATC TCTGGCGGTG  
 3841 TTGACATAAA TACCACTGGC GGTGATACTG AGCACATCAG CAGGACGCAC TGACCACCAT  
 3901 GAAGGTGACG CTCTTAAAAAA TTAAGCCCTG AAGAAGGGCA GCATTCAAAG CAGAAGGCTT  
 3961 TGGGGTGTGT GATACGAAAC GAAGCATTGG GATCATCACA AGTTTGTACA AAAAGCTGA  
 4021 ACGAGAAACG TAAAATGATA TAAATATCAA TATATTAAAT TAGATTTGCA ATAAAAAAACA  
 4081 GACTACATAA TACTGTAAAAA CACAACATAT CCAGTCACTA TGGCGCCGC TAAGTTGGCA  
 4141 GCATCACCCCG ACGCACTTTG CGCCGAATAA ATACCTGTGA CGGAAGATCA CTTCCAGAA  
 4201 TAAATAAAATC CTGGTGTCCC TGTGATACC GGGAAAGCCCT GGGCAACTT TTGGCGAAAA  
 4261 TGAGACGTTG ATCGGCACGT AAGAGGTTCC AACTTTCACC ATAATGAAAT AAGATCACTA  
 4321 CCGGGCGTAT TTTTGAGTT ATCGAGATT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA  
 4381 AAAATCACTG GATATACCCAC CGTTGATATA TCCAATGGC ATCGTAAAGA ACATTTTGAG  
 4441 GCATTCAGT CAGTTGCTCA ATGTACCTAT AACCAAGACCG TTCAGCTGGA TATTACGGCC  
 4501 TTTTTAAAGA CCGTAAAGAA AAATAAGCAC AAGTTTATC CGGCCTTAT TCACATTCTT  
 4561 GCCCCCTGA TGAATGCTCA TCCGGAATTG CGTATGGCAA TGAAAGACGG TGAGCTGGTG  
 4621 ATATGGGATA GTGTTCACCC TTGTTACACC GTTTTCCATG AGCAAACCTGA AACGTTTTCA  
 4681 TCGCTCTGGA GTGAATACCA CGACGATTTC CGGCAGTTTC TACACATATA TTGCAAGAT  
 4741 GTGGCGTGTGTT ACGGTGAAAA CCTGGCTAT TTCCCTAAAG GTTTTATTGA GAATATGTTT  
 4801 TTCGTCAG CCAATCCCTG GGTGAGTTTC ACCAGTTTG ATTTAAACGT GGCAATATG  
 4861 GACAACCTCT TCGCCCCCGT TTTCACCATG GGCACAAATT ATACGCAAGG CGACAAGGTG  
 4921 CTGATGCCGC TGGCGATTCA GTTTCATCAT GCCGCTCTGTG ATGGCTTCCA TGTGGCAGA  
 4981 ATGCTTAATG AATTACAACA GTACTGCGAT GAGTGGCAGG GCGGGCGTA AACCGTGG  
 5041 TCCGGCTTAC TAAAAGCCAG ATAACAGTAT GCGTATTGTC GCGCTGATTT TTGCGGTATA  
 5101 AGAATATATA CTGATATGTA TACCCGAAGT ATGTCAAAAA GAGGTGTGCT ATGAAGCAGC  
 5161 GTATTACAGT GACAGTTGAC AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA  
 5221 TATCTCCGGT CTGGTAAGCA CAACCATGCA GAATGAAGCC CGTCGTCCTGC GTGCCGAACG  
 5281 CTGGAAAGCG GAAAATCAGG AAGGGATGGC TGAGGTCGCC CGGTTTATTG AAATGAACGG  
 5341 CTCTTTGCT GACGAGAACAA GGGACTGGTG AAATGCAGTT TAAGGTTTAC ACCTATAAAA  
 5401 GAGAGAGCCG TTATCGTCTG TTGTTGGATG TACAGAGTGA TATTATTGAC ACGCCGGGC  
 5461 GACGGATGGT GATCCCCCTG GCCAGTGCAC GTCTGCTGTC AGATAAAAGTC TCCCGTGAAC  
 5521 TTTACCCGGT GGTGCATATC GGGGATGAAA GCTGGCGCAT GATGACCAAC GATATGGCCA  
 5581 GTGTGCCGGT CTCCGTTATC GGGGAAGAAG TGCGTGTACT CAGCCACCGC GAAAATGACA  
 5641 TCAAAAACGC CATTAAACCTG ATGTTCTGGG GAATATAAAAT GTCAGGCTCC GTTATACACA  
 5701 GCCAGTCTGC AGGTCGACCA TAGTGAUTGG ATATGTTGTG TTTTACAGTA TTATGTAGTC  
 5761 TGTTTTTAT GCAAAATCTA ATTTAATATA TTGATATTAA TATCATTTA CGTTTCTCGT  
 5821 TCAGCTTCT TGTACAAAGT GGTGATAA

FIGURE 33C

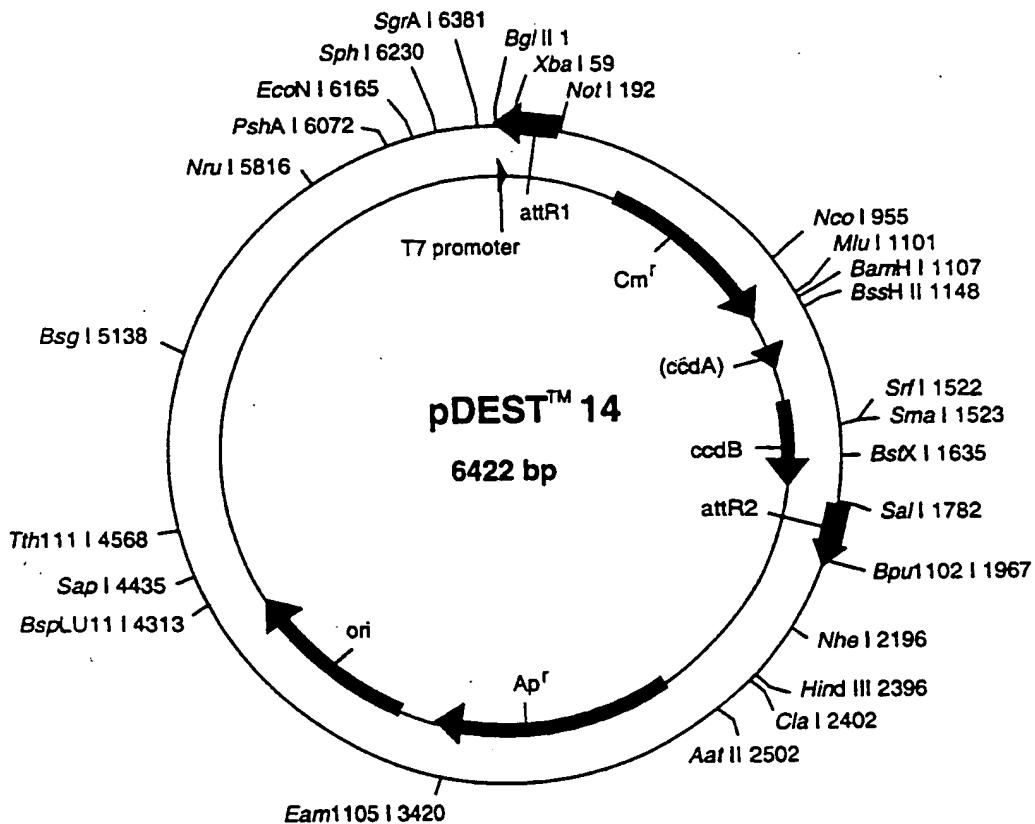
83/240

**Figure 34A: pDEST14 Native Protein Expression in *E. coli*, T7 Promoter**

3961 tgccggccac gatgcgtccg gcgttagagga tcgagatctc gatcccgcga aataataacg  
 acggccgtg ctacgcaggc cgcatttcct agctcttagt ctagggcgct ttaattatgc  
 mEVA

4021 actcaactata gggagaccac aacggtttcc ctctagatca caagtttgta caaaaaaagct  
 tgagtatccatctggtg ttgccaaagg gagatctatgt gttcaaacat gttttttcga

Bgl II      Ase I      PT7 →  
 Xba I      E. coli R1 ↓



84/240

## pDEST14 6422 bp (rotated to position 4000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
185..61	attR1
435..1094	CmR
1214..1298	inactivated ccdA
1436..1741	ccdB
1782..1906	attR2
2632..3489	ampR

1 CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGATC  
 61 ACAAGTTTGT ACAAAAAAGC TGAAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA  
 121 AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA  
 181 CTATGGCGC CGCTAAGTT GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG  
 241 TGACGGAAGA TCACCTCGCA GAATAAATAA ATCTGGTGT CCCTGTTGAT ACCGGGAAGC  
 301 CCTGGGCCAA CTTTGGCGA AAATGAGACG TTGATCGGC CGTAAGAGGT TCCAACTTTC  
 361 ACCATAATGA ATAAGATCA CTACCGGGCG TATTTTTGA GTTATCGAGA TTTTCAGGAG  
 421 CTAAGGAAGC TAAAATGGAG AAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT  
 481 GGCATCGTAA AGAACATTTC GAGGCATTT AGTCAGTTGC TCAATGTACC TATAACCAGA  
 541 CCGITTCAGCT GGATATTACG GCCTTTTAA AGACCGTAA GAAAAATAAG CACAAGTTT  
 601 ATCCGGCTT TATTACATT CTTGCCGCG TGATGAATGC TCATCCGAA TTCCGTATGG  
 661 CAATGAAAGA CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTTGTAC ACCGTTTCC  
 721 ATGAGCAAAC TGAAACGTT TCATCGCTC GGAGTGAATA CCACGACGAT TTCCGGCAGT  
 781 TTCTACACAT ATATTGCAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA  
 841 AAGGGTTTAT TGAGAATATG TTTTCGCTC CAGCCAATCC CTGGGTGAGT TTCACCAAGTT  
 901 TTGATTTAAA CGTGGCAAT ATGGACAATCT TCTTCGCCCC CGTTTCAACC ATGGCAAAT  
 961 ATTATACGCA AGGCGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT  
 1021 GTGATGGCTT CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC  
 1081 AGGGCGGGGC GTAAACGCGT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT  
 1141 TGGCGCTGA TTTTGCGGT ATAAGAATAT ATACTGATAT GTATACCGA AGTATGTCAA  
 1201 AAAGAGGTGT GCTATGAAGC AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT  
 1261 GCTCAAGGC TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAATGAA  
 1321 GCCCGTCGTC TGCCTGCCGAA ACCTGGAAA GCGGAAAATC AGGAAGGGAT GGCTGAGGTC  
 1381 GCCCGGTTTA TTGAAATGAA CGGCTCTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA  
 1441 GTTTAAGGTT TACACCTATA AAAGAGAGAG CGGTTATCGT CTGTTTGTGG ATGTACAGAG  
 1501 TGATATTATT GACACGCCCG GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGCTCTGCT  
 1561 GTCAGATAAA GTCTCCCGTG AACTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG  
 1621 CATGATGACC ACCGATATGG CCAGTGTGCC GGTCCTCGTT ATCGGGGAAG AAGTGGCTGA  
 1681 TCTCAGCCAC CGCGAAAATG ACATAAAAAA CGCATTAAAC CTGATGTTCT GGGGAATATA  
 1741 AATGTCAGGC TCCCTTATAC ACAGCCAGTC TGCAAGTCGA CCATAGTGA TGGATATGTT  
 1801 GTGTTTACA GTATTATGTA GTCTGTTTT TATGCAAAT CTAATTAAAT ATATTGATAT  
 1861 TTATATCATT TTACGTTCT CGITCAGCTT TCTGTACAA AGTGGTGATG ATCCGGCTGC  
 1921 TAACAAAGGC CGAAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA  
 1981 ACCCTTGGG GCCTCTAAAC GGGTCTTGAG GGGTTTTTG CTGAAAGGAG GAACTATATC  
 2041 CGGATATCCA CAGGACGGGT GTGGTCGCCA TGATCGCGTA GTCGATAGTG GCTCCAAGTA  
 2101 GCGAAGCGAG CAGGACTGGG CGGGGGCCAA AGGGTCCGA CAGTGTCTCG AGAACGGGTG  
 2161 CGCATAGAAA TTGCTCAAC GCATATAGCG CTAGCAGCAC GCCATAGTGA CTGGCGATGC  
 2221 TGTCGGAATG GACGATATCC CGCAAGAGGC CGGGCAGTAC CGGCATAACC AAGCCTATGC  
 2281 CTACAGCATC CAGGGTGACG GTGCCGAGGA TGACGATGAG CGCATTTITA GATTTCATAC  
 2341 ACGGTGCCCTG ACTGGCTTAG CAATTTAACT GTGATAAAACT ACCGCAITAA AGCTTATCGA  
 2401 TGATAAGCTG TCAAAACATGA GAATTCTTGA AGACGAAAGG GCCTCGTGTAC AGCCCTATTT  
 2461 TTATAGGTTA ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAG TTTTGGGGA  
 2521 AATGTGCGCG GAACCCCTAT TTGTTTATT TTCTAAATAC ATTCAAATAT GTATCCGTC  
 2581 ATGAGACAAT AACCCCTGATA AATGCTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT  
 2641 CAACATTTCGCTT TATTCCCTTT TTGCGGCAT TTTGCCCTTC TGTTTTGCT  
 2701 ACCCAGAAA CGCTGGTGAAGTAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT-

FIGURE 34B

85/240

2761 TACATCGAAC TGGATCTCAA CAGCGTAAG ATCCTTGAGA GTTTTCGCC CGAAGAACGT  
 2821 TTCCCAATGA TGAGCACTTT TAAAGTCTG CTATGTGGCG CGGTATTATC CCGTGTGAC  
 2881 GCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC  
 2941 TCACCACTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT  
 3001 GCCATAACCA TGAGTGATAA CACTGCGGC AACTTACTTC TGACAACGAT CGGAGGACCG  
 3061 AAGGAGCTAA CCGCTTTTTT GCACAAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG  
 3121 GAACGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA  
 3181 ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCAGACTAC TTACTCTAGC TTCCCAGCAA  
 3241 CAATTAAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT  
 3301 CCGGCTGGCT GTTTTATTGC TGATAAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC  
 3361 ATTGCAGCAC TGCGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG  
 3421 AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT  
 3481 AAGCATTGGT AACTGTCAGA CCAAGTTAC TCATATATAC TTTAGATTGA TTTAAAACCTT  
 3541 CATTTTAAT TAAAAGGAT CTAGGTGAAG ATCCTTTTG ATAATCTCAT GACCAAAATC  
 3601 CCTTAACGTG AGTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT  
 3661 TCTTGAGATC CTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA  
 3721 CCAGGGTGG TTTGTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAACTGGC  
 3781 TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACAC  
 3841 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT  
 3901 GCTGCCAGTG GCGATAAGTC GTGTCTTAC GGGTTGGACT CAAGACGATA GTTACCGGAT  
 3961 AAGGCGCAGC GGTGGGCTG AACGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG  
 4021 ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA  
 4081 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGGAGG  
 4141 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCTG TCGGGTTTCG CCACCTCTGA  
 4201 CTTGAGCGTC GATTTTTGTT ATGCTCGTC GGGGGCGGA GCCTATGGAA AAACGCCAGC  
 4261 AACCGGGCCT TTTTACGGTT CCTGGCCCTT TGCTGGCCTT TTGCTCACAT GTTCTTCCCT  
 4321 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT  
 4381 CGCCGAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCCCTG  
 4441 ATGCGTATT TTCTCCTTAC GCATCTGTG GGTATTTAC ACCGCATATA TGGTGCACCTC  
 4501 TCAGTACAAT CTGCTCTGAT GCGCATAGT TAAGCCAGTA TACACTCCGC TATCGTACG  
 4561 TGACTGGTC ATGGCTCGC CCCGACACCC GCCAACACCC GCTGACGGCC CCTGACGGGC  
 4621 TTGTCGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG  
 4681 TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGGCAG CTGCGTAAAG GCTCATCAGC  
 4741 GTGGCTGTGA AGCGATTAC AGATGTCGTC CTGTTCATCC GCGTCCAGCT CGTTGAGTTT  
 4801 CTCCAGAACG GTTAATGTC GGCTTCTGAT AAAGCGGGCC ATGTTAAGGG CGGTTTTTC  
 4861 CTGTTGGTC ACTGATGCC CCGTGTAAAGG GGGATTCTG TTGATGGGG TAATGATACC  
 4921 GATGAAACGA GAGAGGATGC TCACGATACG GGTTACTGAT GATGAACATG CCCGGTTACT  
 4981 GGAACGTTGT GAGGGTAAAC AACTGGCGGT ATGGATGCCG CGGGACCAAGA GAAAATCAC  
 5041 TCAGGGTCAA TGCCAGCGCT TCGTTAACAC AGATGTTAGGT TTGACACAGG GTAGCCAGCA  
 5101 GCATCCTCGC ATGCAGATCC GGAACATAAT GGTGCAGGGC GCTGACTTCC GCGTTTCCAG  
 5161 ACTTTACGAA ACACGGAAAC CGAACGACCAT TCATGTTGTT GTCAGGTCG CAGACGTTT  
 5221 GCAGCAGCAG TCGCTTCACG TTGCTCGCG TATCGGTGAT TCATCTGCT AACCAAGTAAG  
 5281 GCAACCCCGC CAGCTTAGCC GGGTCTCAA CGACAGGAGC AGCATCATGC GCACCCGTGG  
 5341 CCAGGACCCA ACCTGCCCC AGATGCCCG CGTGGCGCTG CTGGAGATGG CGGACCGAT  
 5401 GGATATGTTG TCCAAGGGT TGGTTGCGC ATTACAGTT CTCCGCAAGA ATTGATTGGC  
 5461 TCCAATTCTT GGAGTGGTGA ATCCGTTAGC GAGGTGCCG CGGCTTCCAT TCAGGTCGAG  
 5521 GTGGCCGGC TCCATGACCG GCGACCCAC GCGGGGAGGC AGACAAGGTA TAGGGCGGG  
 5581 CCTAACATCC ATGCCAACCC GTTCCATGTG CTCCCGAGG CGGCATAAAT CGCCGTGACG  
 5641 ATCAGCGGTC CAGTGATCGA AGTTAGGCTG GTAAGAGCG CGAGCGATCC TTGAAGCTGT  
 5701 CCCTGATGGT CGTCATCTAC CTGCGTGGAC AGCATGCCCT GCAACGCGGG CATCCCGATG  
 5761 CGGCCGGAAG CGAGAAGAAT CATAATGGG AAGGCCATCC AGCCTCGCGT CGCGAACGCC  
 5821 AGCAAGACGT AGCCCAGCGC GTCCGGGCC ATGCCGGCA TAATGGCTG CTTCTCGCCG  
 5881 AAACCTTGG TGGCGGGACC AGTGACCAAG GCTTGAGCGA GGGCGTGCAGA GATTCCGAAT  
 5941 ACCGCAAGCG ACAGGCCGAT CATGTCGCG CTCCAGCGAA AGCGGTCCTC GCGAAAATG  
 6001 ACCCAGAGCG CTGCCGGCAC CTGCTCTACG AGTTGATCGA TAAAGAAGAC AGTCATAAGT  
 6061 GCGGCAGCA TAGTCATGCC CCGCGCCAC CGGAAGGAGC TGACTGGGTT GAAGGCTCTC  
 6121 AAGGGCATCG GTCGATCGAC GCTCTCCCTT ATGCGACTCC TGCAATTAGGA AGCAGCCAG  
 6181 TAGTAGGTTG AGGCCGTTGA GCACCGCCGC CGCAAGGAAT GGTGCATGCA AGGAGATGGC-

FIGURE 34C

86/240

6241 GCCCAACAGT CCCCCGGCCA CGGGGCCTGC CACCATAACCC ACGCCGAAAC AAGCGCTCAT  
6301 GAGCCCGAAG TGGCGAGCCC GATCTTCCCC ATCGGTGATG TCGGCGATAT AGGCGCCAGC  
6361 AACCGCACCT GTGGCGCCGG TGATGCCGGC CACCGATGCGT CCGGCGTAGA GGATCGAGAT  
6421 CT

FIGURE 34D

87/240

**Figure 35A: pDEST15 Glutathione-S-transferase Fusion in *E. coli*, T7 Promoter**

mRNA

T7 Promoter

1 nat cga gat ctc gat ccc gcg aaa tta ata cga ctc act ata |ggg| aga cca  
 nta gct cta gag cta ggg cgc ttt dat tat gct gag tga tat ccc tct ggt

52 caa cgg ttt ccc tat aga aat aat ttt gtt taa ctt taa gaa gga gat ata  
 gtt gcc aaa ggg aga tat tta aaa caa att gaa att ctt cct cta tat

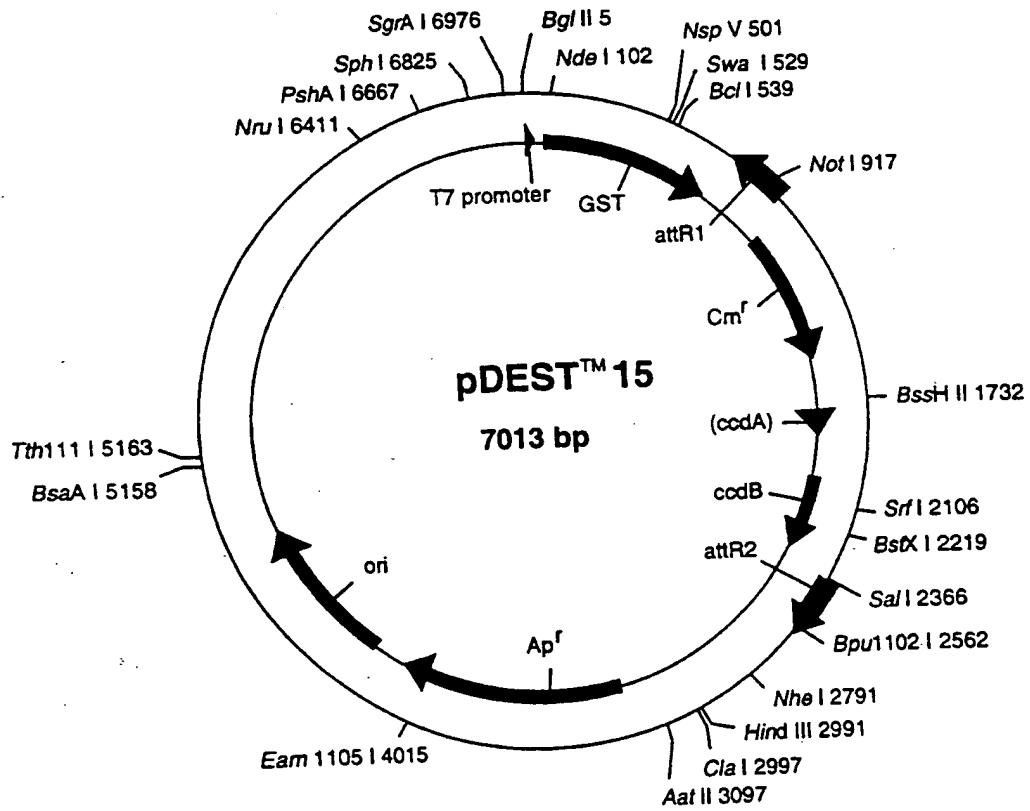
103 XbaI NdeI A S P I L T7 Promoter GST  
 cat atg tcc cct ata cta ggt tat tgg aaa att aag ggc ctt gtg caa ccc  
 gta tac agg gga tat gat cca ata acc ttt taa ttc ccc gaa cac gtt ggg

154 Start Translation GST  
 act cga ctt ctt ttg gaa tat ctt gaa gaa aaa tat gaa gag 'cat ttg tat  
 tga gct gaa gaa aac ctt ata gaa ctt ctt ttt ata ctt ctc gta aac ata

715 cag ggc tgg caa gcc acg ttt ggt ggt ggc gac cat cct cca aaa tcg gat  
 gtc ccc acc gtt cgg tgc aaa cca cca ccc ctg gta gga ggt ttt agc cta

766 ctg gtt ccc cgt cca tgg tgg aat cca attR1 aca agt tgg tac aaa aaa gct gaa  
 gac cca ggc gca ggt acc agc tta gtt tgt tca aac atg ttt ttt cga ctt //

817 cga gaa acg taa aat gat ata aat atc aat ata tta aat tag att ttg cat  
 gct ctt tgc att tta cta tat tta tag tta tat aat tta atc taa aac gta



88/240

## pDEST15 7013 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
108..776	GST
916..792	attR1
1025..1537	CmR
1804..1888	inactivated ccdA
2026..2331	ccdB
2372..2496	attR2
3233..4093	ampR

1 ATCGAGATCT CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC  
 61 CCTCTAGAAA TAATTGTT TAACCTTAAAG AAGGAGATAT ACATATGTCC CCTATACTAG  
 121 GTTATTGGAA AATTAAGGGC CTTGTGCAAC CCACCTGACT TCTTTGGAA TATCTTGAAG  
 181 AAAAATATGA AGAGCATTTG TATGAGCGCG ATGAAGGTGA TAAATGGCGA AACAAAAAGT  
 241 TTGAATTGGG TTTGGAGTTT CCCAACCTTC CTTATTATAT TGATGGTGT GTTAAATTAA  
 301 CACAGTCTAT GCCCATCATA CGTTATATAG CTGACAAGCA CAACATGTG GGTGGTTGTC  
 361 CAAAAGAGCG TGCAGAGATT TCAATGCTTG AAGGAGCGGT TTTGGATATT AGATACGGTG  
 421 TTTCGAGAAT TGCATATAGT AAAGACTTTG AAACCTCTAA AGTTGATTTT CTTAGCAAGC  
 481 TACCTGAAAT GCTGAAATG TTCGAAGATC GTTTATGTCA TAAAACATAT TTAAATGGTG  
 541 ATCATGTAAC CCATCCTGAC TTCAATGTTGT ATGACGCTCT TGATGTTGT TTATACATGG  
 601 ACCCAATGTC CCTGGATGCG TTCCCCAAAT TAGTTGTT TAAAAAAACGT ATTGAAGCTA  
 661 TCCCACAAAT TGATAAGTAC TTGAAATCCA GCAAGTATAT AGCATGGCCT TTGCAGGGCT  
 721 GGCAAGCCAC GTTTGGTGGT GGCGACCATC CTCCAAAATC GGATCTGGTT CCGCGTCCAT  
 781 GGTCGAATCA AACAAAGTTG TACAAAAAAG CTGAACGAGA AACGTAAT GATATAAATA  
 841 TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC  
 901 ATATCCAGTC ACTATGGCGG CCGCATTAGG CACCCCAGGC TTACACTTT ATGCTTCCGG  
 961 CTCGTATAAT GTGTGGATTT TGAGTTAGGA TCCGTCGAGA TTTTCAGGAG CTAAGGAAGC  
 1021 TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAAT GGCATCGTAA  
 1081 AGAACATTTT GAGGCATTTC AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAGCT  
 1141 GGATATTACG GCCTTTTAA AGACCGTAA GAAAAATAAG CACAAGTTTT ATCCGGCCTT  
 1201 TATTCACTATT CTTGCCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA  
 1261 CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTTGTAC ACCGTTTTCC ATGAGCAAAC  
 1321 TGAAACGTTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT  
 1381 ATATTGCAA GATGTGGCGT GTTAACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT  
 1441 TGAGAATATG TTTTCGCTCT CAGCCAATCC CTGGGTGAGT TTCACCACTT TTGATTAAA  
 1501 CGTGGCAAT ATGGACAAC TCTTCGCCCC CGTTTCAAC ATGGGCAAAT ATTATACGCA  
 1561 AGGCAGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT  
 1621 CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGC  
 1681 GTAATCTAGA GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TGCGCGTGA  
 1741 TTTTGCGGT ATAAGAATAT ATACTGATAT GTATACCGA AGTATGTCAA AAAGAGGTGT  
 1801 GCTATGAAGC AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA  
 1861 TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAATGAA GCCCCTCGTC  
 1921 TGCCTGCCGA ACGCTGGAAA GCGGAAATC AGGAAGGGAT GGCTGAGGTG GCCCCTTTA  
 1981 TTGAAATGAA CGGCTCTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTAAAGGTT  
 2041 TACACCTATA AAAGAGAGAG CGCTTATCGT CTGTTGTGG ATGTACAGAG TGATATTATT  
 2101 GACACGCCCG GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGTCTGCT GTCAGATAAA  
 2161 GTCTCCCGTG AACTTTACCC GGTGGTGCAT ATCAGGGATG AAAGCTGGCG CATGATGACC  
 2221 ACCGATATGG CCAGTGTGCC GGTCTCCGTT ATCAGGGAG AAGTGGCTGA TCTCAGGCCAC  
 2281 CGCGAAATG ACATCAAAA CGCCATTAAC CTGATGTTCT GGGGAATATA AATGTCAGGC  
 2341 TCCCTTATAC ACAGCCAGTC TGCAAGTCGA CCATAGTGAC TGGATATGTT GTGTTTACA  
 2401 GTATTATGTA GTCTGTTTT TATGCAAAAT CTAATTAAAT ATATTGATAT TTATATCATT  
 2461 TTACGTTCT CGTTCAGCTT TCTTGTACAA AGTGGTTGATCCTGACCCGG GATCCGGCTG  
 2521 CTAACAAAGC CCGAAAGGAA GCTGAGTTGG CTGCTGCCAC CGCTGAGCAA TAACTAGCAT  
 2581 AACCCCTTGG GGCTCTAA CGGGTCTTGA GGGGTTTTT GCTGAAAGGA GGAACATATAT  
 2641 CGGGATATCC ACAGGACGGG TGTGGTCGCC ATGATCGCGT AGTCGATAGT GGCTCCAAGT-

FIGURE 35B

89/240

2701 AGCGAAGCGA GCAGGACTGG GCGGCCGCCA AAGCGGTGG ACAGTGCTCC GAGAACGGGT  
 2761 GCGCATAGAA ATTGCATCAA CGCATATAGC GCTAGCAGCA CGCCATAGTG ACTGGCGATG  
 2821 CTGTCGAAT GGACGATATC CCGCAAGAGG CCCGGCAGTA CGGCATAAC CAAGCCTATG  
 2881 CCTACAGCAT CCAGGGTGAC GGTGCCGAGG ATGACGATGA GCGCATTGTT AGATTCATA  
 2941 CACGGTGCCT GACTGCGTTA GCAATTAAAC TGTGATAAAC TACCGCATT AAGCTTATCG  
 3001 ATGATAAGCT GTCAAACATG AGAATTCTTG AAGACGAAAG GGCCTGTGA TACGCCTATT  
 3061 TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG  
 3121 AAATGTGCGC GGAACCCCTA TTGTTTATT TTCTCAAATA CATTCAAATA TGATCCGCT  
 3181 CATGAGACAA TAACCCGTAT AAATGCTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT  
 3241 TCAACATTC CGTGTGCCCT TTATTCCCTT TTTTGCAGCA TTTTGCCTTC CTGTTTTG  
 3301 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG  
 3361 TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG  
 3421 TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTGTTGA  
 3481 CGCCGGGCAA GAGCAACTCG GTCGCCCAT ACACATTCT CAGAATGACT TGGTTGAGTA  
 3541 CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAAT TATGCAGTGC  
 3601 TGCCATAACC ATGAGTGTATA ACAC TGCGGC CAACTTACTT CTGACAAACGA TCGGAGGACC  
 3661 GAAGGAGCTA ACCGCTTTTT TGACAAACAT GGGGGATCAT GTAACCGGCC TTGATCGTTG  
 3721 GGAACCGGAG CTGAATGAAG CCATACAAAA CGACGAGCGT GACACCACGA TGCCTGCAGC  
 3781 AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGCA  
 3841 ACAATTAAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACCTCTGC GCTCGCCCT  
 3901 TCCGGCTGGC TGTTTATTG CTGATAAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT  
 3961 CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCGTATC CTAGTTATCT ACACGACGGG  
 4021 GAGTCAGGCA ACTATGGATG AACGAAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT  
 4081 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAATC  
 4141 TCATTTTAA TTAAAAAGGA TCTAGGTGA GATCCTTTT GATAATCTCA TGACCAAAAT  
 4201 CCCTTAACGT GAGTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC  
 4261 TTCTTGAGAT CCTTTTTTTTC TGCGCGTAAT CTGCGCTTG CAAACAAAAA AACCACCGCT  
 4321 ACCAGCGGTG GTTGTGTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACTGG  
 4381 CTTCAAGAAC GCGCAGATAC CAAACTGT CTTCTAGTG TAGCGTAGT TAGGCCACCA  
 4441 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCTGT TACCAAGTGGC  
 4501 TGCTGCCAGT GCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA  
 4561 TAAGGCGCAG CGGTGGGGCT GAACGGGGGG TTCTGTCACA CAGCCAGCT TGGAGCGAAC  
 4621 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCGA  
 4681 AGGGAGAAAG GCGGACAGGT ATCCGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG  
 4741 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCTGGTTTC GCCACCTCTG  
 4801 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGCGG AGCCTATGGA AAAACGCCAG  
 4861 CAACCGGCC TTGTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC  
 4921 TGCCTTATCC CCTGATTCTG TGGATAACCG TATTACGCC TTTGAGTGAG CTGATACCGC  
 4981 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCT  
 5041 GATGCCGTAT TTCTCCTTA CGCATCTGT CGGTATTTC CACCGCATAT ATGGTGCACT  
 5101 CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGT ATACACTCCG CTATCGCTAC  
 5161 GTGACTGGGT CATGGCTGCC CCCCCACACC CGCCAACACC CGCTGACCGCG CCCTGACGGG  
 5221 CCTGCTGCT CCGGCATCC GCTTACAGAC AAGCTGTGAC CGTCTCCGGG AGCTGCATGT  
 5281 GTCAGAGGTT TTCACCGTCA TCACCGAAC GCGCGAGGCA GCTGCGTAA AGCTCATCAG  
 5341 CGTGGTCGTG AAGCGATTCA CAGATGTCTG CCTGTTCATC CGCGTCCAGC TCGTTGAGTT  
 5401 TCTCCAGAAC CGTTAATGTC TGGCTCTGA TAAAGCGGGC CATGTTAAAGG GCGGTTTTT  
 5461 CCTGTTGGT CACTGATGCC TCCGTTAAG GGGGATTCT GTTCATGGGG GTAATGATAC  
 5521 CGATGAAACG AGAGAGGATG CTCACGATAC GGGTTACTGA TGATGAACAT GCCCGGTTAC  
 5581 TGGAACGTTG TGAGGGTAAA CAACTGGCGG TATGGATGCG GCGGGACCAAG AGAAAAAATCA  
 5641 CTCAGGGTCA ATGCCAGGCC TTGCTTAATA CAGATGTAGG TGTTCCACAG GGTAGCCAGC  
 5701 AGCATCCTGC GATGCAGATC CGGAACATAA TGGTGCAGGG CGCTGACTTC CGCGTTCCA  
 5761 GACTTTACGA AACACGGAAA CGGAAGACCA TTCATGTTGT TGCTCAGGTC GCAGACGTTT  
 5821 TGCAGCAGCA GTCGCTTCAC GTTCGCTCGC GTATCGGTGA TTCAATTCTGC TAACCAGTAA  
 5881 GGCAACCCCG CCAGCCTAGC CGGGTCCCTCA ACGACAGGAG CACGATCATG CGCACCCGTG  
 5941 GCCAGGACCC AACGCTGCC GAGATGCC GCGTCCGGCT GCTGGAGATG GCGGACCGA  
 6001 TGGATATGTT CTGCCAAGGG TTGGTTGCG CATTACAGT TCTCCGCAAG AATTGATTGG  
 6061 CTCCAATTCT TGGAGTGGTG AATCCGTTAG CGAGGTGCCG CCGGCTTCCA TTCAGGTGCA  
 6121 GGTGGCCCGG CTCCATGCA CGCGACGCAA CGGGGGGAGG CAGACAAGGT ATAGGGCGGC-

FIGURE 35C

90/240

6181 GCCTACAATC CATGCCAACCG CGTTCCATGT GCTCGCCGAG GCGGCATAAA TCGCCGTGAC  
6241 GATCAGCGGT CCAGTGATCG AAGTTAGGCT GGTAAGAGCC GCGAGCGATC CTTGAAGCTG  
6301 TCCCTGATGG TCGTCATCTA CCTGCCTGGA CAGCATGGCC TCACAACGCGG GCATCCCGAT  
6361 GCCGCCGGAA GCGAGAAGAA TCATAATGGG GAAGGCCATC CAGCCTCGCG TCGCGAACGC  
6421 CAGCAAGACG TAGCCCAGCG CGTCGGCCGC CATGCCGGCG ATAATGGCCT GCTTCTCGCC  
6481 GAAACGTTTG GTGGCGGGAC CAGTGACGAA GGCTTGAGCG AGGGCGTGCA AGATTCCGA  
6541 TACCGCAAGC GACAGGCCGA TCATCGTCGC GCTCCAGCGA AAGCGGTCT CGCCGAAAAAT  
6601 GACCCAGAGC GCTGCCGGCA CCTGTCTTAC GAGTTGCATG ATAAAGAAGA CAGTCATAAG  
6661 TGCGGCGACG ATAGTCATGC CCCGCGCCCA CCGGAAGGAG CTGACTGGGT TGAAGGCTCT  
6721 CAAGGGCATC GGTGATCGA CGCTCTCCCT TATGCGACTC CTGCATTAGG AAGCAGCCCCA  
6781 GTAGTAGGTT GAGGCCGTTG AGCACCGCCG CCGCAAGGAA TGGTGCATGC AAGGAGATGG  
6841 CGCCCAACAG TCCCCCGGCC ACGGGGCCTG CCACCATACC CACGCCGAAA CAAGCGCTCA  
6901 TGAGCCCGAA GTGGCGAGCC CGATCTTCCC CATCGGTGAT GTCGGCGATA TAGGCGCCAG  
6961 CAACCGCACC TGTGGCGCCG GTGATGCCGG CCACCGATGCG TCCGGCGTAG AGG

FIGURE 351)

91/260

Figure 36A:  $\lambda$ DEST16Thioredoxin N-Fusion Protein  
in E. coli with T7 Promoter

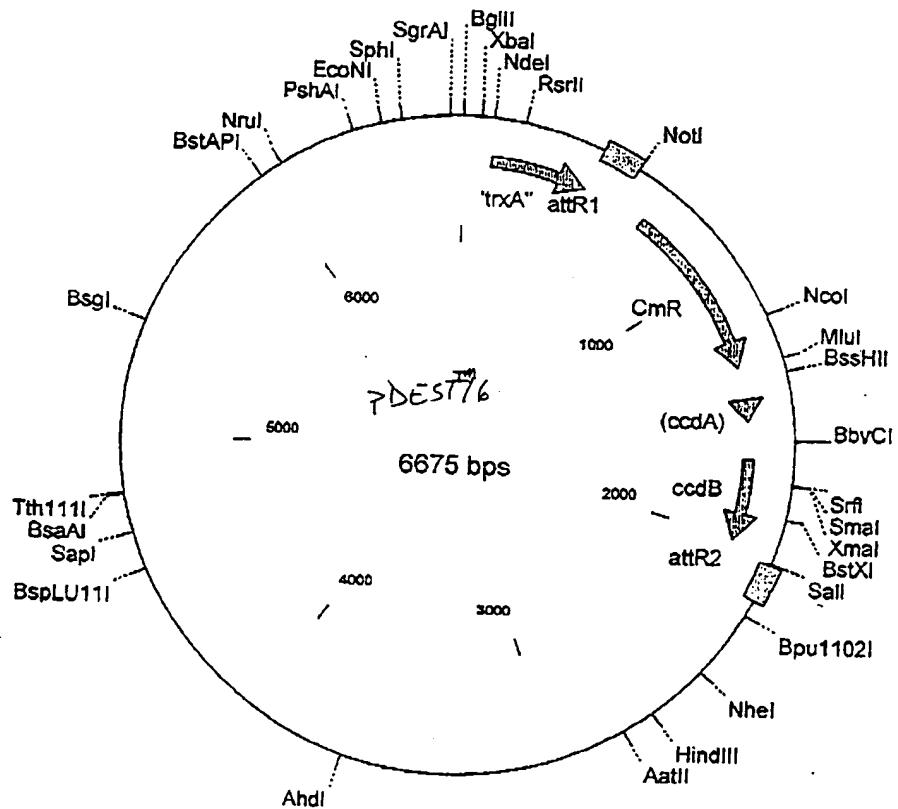
mRNA

T7 Promoter

```

1  gat ctc gat ccc gcg aaa tca ata cga ctc act ata ggg aga cca caa cgg
   cta gag cta ggg cgc ttt aat tat gct gag tga tat ccc tct ggt gtt gcc
   XbaI
52  ttt ccc tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata cat atg Start
   aaa ggg aga tct tta tta aaa caa att gaa att ctt cct cta tat gta tac Translation Trx
103  S D K - - - agc gat aaa att att cac ctg act gac gac agt ttt gac acg gat gta ctc
   tcg cta ttt taa taa gtg gac tga ctg ctg tca aaa ctg tgc cta cat gag
//358  gtg gcg gca acc aaa gtg ggt gca ctg tct aaa ggt cag ttg aaa gag ttc
   cac cgc cgt tgg ttt cac cca cgt gac aga ttt cca gtc aac ttt ctc aag
409  ctc gac gct aac ctg gcc ggt tct ggt tct ggt gat gac gat gac aag atc
   gag ctg cga ttg gac cgg cca aga cca aga cca cta ctg cta ctg ttc tag
   T S L Y K K A attR1
460  aca agt ttg tac aaa aaa gct gaa cga gaa acg taa aat gat ata aat atc
   ttt cga ctt gct ctt tgc att tta cta tat tta tag
   IntV

```



92/240

## pDEST16 6675 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
104..457	trxA
585..461	attR1
694..1353	CmR
1473..1557	inactivated ccdA
1695..2000	ccdB
2041..2165	attR2
1 AGATCTCGAT CCCCGCAAAT TAATACGACT CACTATAGGG AGACCACAAC GGTTTCCCTC	
61 TAGAAATAAT TTGTTTAAC TTTAAGAAGG AGATATACAT ATGAGCGATA AAATTATTCA	
121 CCTGACTGAC GACAGTTTG ACACGGATGT ACTCAAAGCG GACGGGGCGA TCCTCGTCGA	
181 TTTCTGGCA GAGTGGTGCG GTCCGTGCAA AATGATGCC CCGATTCTGG ATGAAATCGC	
241 TGACGAATAT CAGGGCAAAC TGACCGTGC AAAACTGAAC ATCGATCAA ACCCTGGCAC	
301 TGCGCCAAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA ACGGTGAAGT	
361 GGC GGCAACC AAAGTGGGTG CACTGCTAA AGGTCAAGTTG AAAGAGTTCC TCGACGCTAA	
421 CCTGGCCGGT TCTGGTTCTG GTGATGACGA TGACAAAGATC ACAAGTTTGT ACAAAAAAAGC	
481 TGAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA ATTAGATTG TGCATAAAAAA	
541 ACAGACTACA TAATACTGTA AAACACAAACA TATCCAGTCA CTATGGCGGC CGCATTAGGC	
601 ACCCCAGGCT TTACACTTTA TGCTTCCGGC TCGTATAATG TGTGGATTG GAGTTAGGAT	
661 CCGGCGAGAT TTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAAATCAC TGGATATACC	
721 ACCGTTGATA TATCCCAATG GCATCGTAA GAACATTG AGGCATTCA GTCAGTTGCT	
781 CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG CCTTTTTAAA GACCGTAAAG	
841 AAAAATAAGC ACAAGTTTTA TCCGGCCTT ATTACACATTG TTGCCCCGCT GATGAATGCT	
901 CATCCGGAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA TAGTGTTCAC	
961 CCTTGTACA CCGTTTCCA TGAGCAAATC GAAACGTTT CATCGCTCTG GAGTGAATAC	
1021 CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG ATGTGGCGTG TTACGGTGAA	
1081 AACCTGGCCT ATTTCCCTAA AGGGTTTATT GAGAATATGT TTTTCGTCTC AGCCAATCCC	
1141 TGGGTGAGTT TCACCAAGTT TGATTTAAC GTGGCCAATA TGGACAACCTT CTTGCC	
1201 GTTTTACCA TGGGCAAATA TTATACGAA GGCACAAAGG TGCTGATGCC GCTGGCGATT	
1261 CAGGTTCATC ATGCCGTCTG TGATGGCTTC CATGTCGGCA GAATGCTTAA TGAATTACAA	
1321 CAGTACTGCG ATGAGTGGCA GGGCGGGCG TAAACGCGTG GATCCGGCTT ACTAAAAGCC	
1381 AGATAAACAGT ATGCGTATTT GCGCGTGTAT TTTTGCCTGA TAAGAATATA TACTGATATG	
1441 TATACCCGAA GTATGTCAAA AAGAGGTG TGATGAAGCA GCGTAAITACA GTGACAGTTG	
1501 ACAGCGACAG CTATCAGTTG CTCAAGGCAT ATATGATGTC AATATCTCCG GTCTGGTAAG	
1561 CACAACCATG CAGAATGAAG CCCGTCGTCT GCGTCCGAA CCCTGGAAAG CGGAAAATCA	
1621 GGAAGGGATG GCTGAGGTGCG CCCGGTTAT TGAAATGAAC GGCTCTTTTG CTGACGAGAA	
1681 CAGGGACTGG TGAAAATGCAG TTAAAGGTTT ACACCTATAA AAGAGAGAGC CGTTATCGTC	
1741 TGTGGTGGG TGAGTACAGAGT GATATTATG ACACGCCCGG GCGACGGATG GTGATCCCC	
1801 TGGCCAGTGC ACCTCTGCTG TCAGATAAAG TCTCCCGTGA ACTTTACCCG GTGGTGCATA	
1861 TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC CAGTGTGCCG GTCTCCGTTA	
1921 TCGGGGAAGA AGTGGCTGAT CTCAGCCACC GCGAAAATGA CATAAAAAAC GCCATTAACC	
1981 TGATGTTCTG GGGAAATAAA ATGTCAGGCT CCCTTATACA CAGCCAGTCT GCAGGTCGAC	
2041 CATAGTGAATGGGATGTTG TGTTTACAG TATTATGTAG TCTGTTTTTG ATGAAAATC	
2101 TAATTTAATA TATTGATATT TATATCATTG TACGTTCTC GTTCAGCTTT CTTGTACAAA	
2161 GTGGTGTGATG TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG	
2221 CTGAGCAATA ACTAGCATAA CCCCTGGGG CCTCTAAACG GGTCTTGAGG GGTTTTTG	
2281 TGAAAGGAGG AACTATATCC GGATATCCAC AGGACGGGTG TGGTCGCCAT GATCGCGTAG	
2341 TCGATAGTGG CTCCAAGTAG CGAAGCGAGC AGGACTGGGC GCGGGCCAAA GCGGTCGGAC	
2401 AGTGCTCCGA GAACGGGTGC GCATAGAAAT TGCACTAACG CATATAGCGC TAGCAGCACG	
2461 CCATAGTGAC TGGCGATGCT GTCGGAATGG ACGATATCCC GCAAGAGGCC CGGCAGTACC	
2521 GGCATAACCA AGCCTATGCC TACAGCATCC AGGGTGCAGG TGCCGAGGAT GACGATGAGC	
2581 GCATTGTTAG ATTTCATACA CGGTGCCTGA CTGCGTTAGC AATTTAACTG TGATAAAACTA	
2641 CCGCATTAAA GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTTGAA GACGAAAGGG	
2701 CCTCGTGATA CGCTTATTT TATAGGTTAA TGTCAATGATA ATAATGGTTT CTTAGACGTC	
2761 AGGTGGCACT TTTCGGGAA ATGTCGCGGG AACCCCTATT TGTTTATTT TCTAAATACA-	

FIGURE 36B

93/240

2821 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA  
 2881 AAGGAAGAGT ATGAGTATTTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT  
 2941 TTGCTTCCTT GTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA  
 3001 GTTGGGTGCA CGAGTGGGTT ACATCGAAC GGATCTCAAC AGCGGTAAGA TCCTTGAGAG  
 3061 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC  
 3121 GGTATTATCC CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA  
 3181 GAATGACTTG GTTGGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT  
 3241 AAGAGAATTG TGCAGTGCTG CCATAACCAC GAGTGATAAC ACTGCGGCCA ACTTACTTCT  
 3301 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT  
 3361 AACTCGCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA  
 3421 CACCAACGATC CCTGCAGCAA TGGCAACAAAC GTTGCAGCAA CTATTAACATG GCGAAGACTACT  
 3481 TACTCTAGCT TCCCAGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGGACC  
 3541 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTATTGCT GATAAATCTG GAGCCGGTGA  
 3601 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT  
 3661 AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA  
 3721 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT  
 3781 TTAGATTGAT TTAAAACCTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTG  
 3841 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTCGTTT CACTGAGCGT CAGACCCCGT  
 3901 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA  
 3961 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCGC GATCAAGAGC TACCAACTCT  
 4021 TTTTCCGAAG GTAACTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA  
 4081 GCCGTAGTTA GGCCACCACT TCAAGAACTC TGAGCACCG CCTACATACC TCGCTCTGCT  
 4141 AATCTGTAA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGAACCTC  
 4201 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCCGGCTGA ACGGGGGGTT CGTGCACACA  
 4261 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGT AGCTATGAGA  
 4321 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CGGTAAAGCG GCAGGGTCGG  
 4381 AACAGGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTT ATAGTCCCTGT  
 4441 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG  
 4501 CCTATGGAAA AACGCCAGCA ACGCGCCCTT TTACGGTTT CTGGCCTTTT GCTGGCCTTIT  
 4561 TGCTCACATG TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT  
 4621 TGAGTGAGCT GATAACGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA  
 4681 GGAAGCGGAA GAGCGCCTGA TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTCACA  
 4741 CCGCATATAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT  
 4801 ACACCTCGCT ATCGCTACGT GACTGGTCA TGGCTGCGCC CCGACACCCG CCAACACCCG  
 4861 CTGACCGGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG  
 4921 TCTCCGGAG CTGCATGTGT CAGAGTTTT CACCGTCATC ACCGAAACCG GCGAGGCAGC  
 4981 TGCCTGAAAG CTCATCAGCG TGGCTGTGAA GCGATTACAA GATGTCTGCC TGTTCATCCG  
 5041 CGTCCAGCTC GTTGGAGTTTC TCCAGAACG TTAATGTCTG GCTTCTGATA AAGCAGGGCCA  
 5101 TGTTAAGGGC GGTTTTTCTC TGTTGGTCA CTGATGCCCTC CGTGTAAAGGG GGATTCTGT  
 5161 TCATGGGGGT AATGATACCG ATGAAACGAG AGAGGATGCT CACGATACGG GTTACTGTATG  
 5221 ATGAACATGC CCGGTTACTG GAACGTTGT AGGTAAACAA ACTGGCGGT AGGATGCGGC  
 5281 GGGACACAGAG AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAAATACA GATGTAGGTG  
 5341 TTCCACAGGG TAGCCAGCAG CATCCTCGCA TGCGAGATCCG GAACATAATG GTGCAGGGCG  
 5401 CTGACTTCCG CGTTTCCAGA CTTTACGAAA CACGGAAACCG GAAGACCATT CATGTTGTTG  
 5461 CTCAGGTGCG AGACGTTTTG CAGCAGCGAGT CGCTTCACGT TCGCTCGCGT ATCGGTGATT  
 5521 CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCAAC GACAGGAGCA  
 5581 CGATCATCGC CACCCGTGGC CAGGACCCAA CGCTGCCGA GATGCGCCGC GTGCGGCTGC  
 5641 TGGAGATGGC GGACCGCAGT GATATGTTCT GCCAAGGGTT GTTTTGCAGCA TTCACAGTT  
 5701 TCCGCAAGAA TTGATTGGCT CCAATTCTG GAGTGGTGA TCCGTTAGCG AGGTGCCGCC  
 5761 GGCTTCCATT CAGGTGCGAGG TGGCCCGCT CCATGCACCG CGACGCAACG CGGGGAGGCA  
 5821 GACAAGGTAT AGGGCGGCCG CTACAATCCA TGCCAAACCG TTCCATGTGC TCGCCGAGGC  
 5881 GGCATAAAATC GCGGTGACGA TCAGCGGTCC AGTGTGAA GTTAGGCTGG TAAGAGCCGC  
 5941 GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTACC TGCCTGGACA GCATGGCCTG  
 6001 CAACCGGGGC ATCCCGATGC CGCCGGAAAGC GAGAAGAATC ATAATGGGAA AGGCCATCCA  
 6061 GCCTCGCGTC GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGCCGCCA TGCCGGCGAT  
 6121 AATGGCCTGC TTCTCGCCGA AACGTTGGT GGCGGGACCA GTGACGAAGG CTTGAGCGAG  
 6181 GGCCTGCAAG ATTCCGAATA CCGCAAGCGA CAGGCCGATC ATCGCTCGCG TCCAGCGAAA  
 6241 GCGGTCTCG CGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCTACGA GTTGCGAT-

FIGURE 36C

94/240

6301 AAAGAAGACA GTCATAAGTG CGCGGACGAT AGTCATGCC CGCGCCCACC GGAAGGGAGCT  
6361 GACTGGGTTG AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TGCGACTCCT  
6421 GCATTAGGAA GCAGCCCAGT AGTAGGTTGA GGCGGTTGAG CACCGCCGCC GCAAGGAATG  
6481 GTGCATGCAA GGAGATGGCG CCCAACAGTC CCCCGGCCAC GGGGCCTGCC ACCATAACCA  
6541 CGCCGAAACA AGCGCTCATG AGCCCCAAGT GGCAGCCCCG ATCTTCCCCA TCGGTGATGT  
6601 CGGGATATA GGCGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC  
6661 CGGCGTAGAG GATCG

FIGURE 36D

95/240

mRNA

T7 Promoter

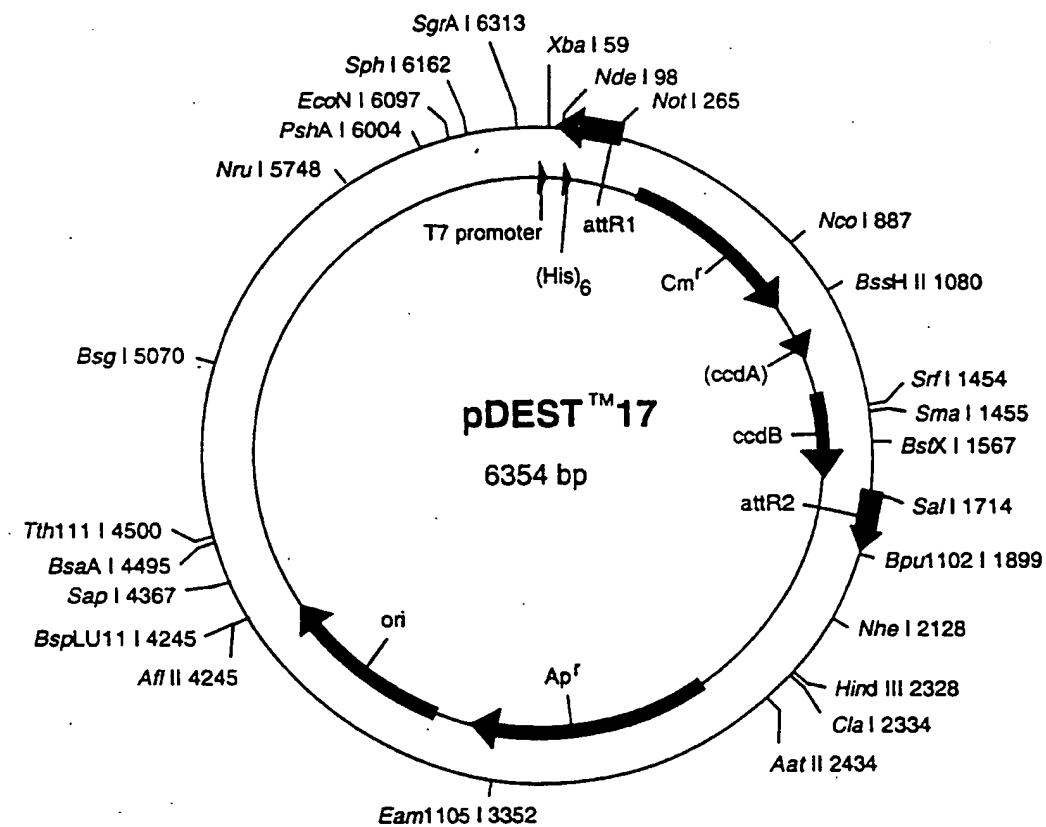
Start Translation

attR1      Int

```

1  gat ccc gcg aaa tta ata cga ctc act ata ggg aga cca caa cgg ttt ccc
   cta ggg cgc ttt aat tat get gag tga tat ccc tct ggt gtt gcc aaa ggg
52  tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata cat atg tcg tac
   aga tct tta tta aaa caa att gaa att ctt cct cta tat gta tac agc atg
103 H H H H L E S T S L Y K K A
   tac cat cac cat cac cat cac ctc gaa tca aca agt tcg tac aaa aaa gct
   atg gta gtg gta gtg gta gtg gag ctt agt tgt tca aac atg ttt ttt cga

```



96/240

## pDEST17 6354 bp

Location (Base Nos.)      Gene Encoded

258..134	attR1
367..1026	CmR
1146..1230	inactivated ccdA
1368..1673	ccdB
1714..1838	attR2
2564..3421	ampR

1 CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGAAA  
 61 TAATTTGTT TAACTTAAAG AAGGAGATAT ACATATGTCG TACTACCATC ACCATCACCA  
 121 TCACCTCGAA TCAACAAGTT TGTACAAAAA AGCTAACGAA GAAACGTAAA ATGATATAAA  
 181 TATCAATATA TAAATTAGA TTTGCATAA AAAACAGACT ACATAATACT GTAAAACACA  
 241 ACATATCCAG TCACTATGGC GGCGCATTA GGCACCCCCAG GCTTTACACT TTATGCTTCC  
 301 GGCTCGTATA ATGTGTGGAT TTTGAGTTAG GATCCGTCGA GATTTTCAGG AGCTAAGGAA  
 361 GCTAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCC ATGGCATCGT  
 421 AAAGAACATT TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCCTTCAG  
 481 CTGGATATTA CGGCCCTTTT AAAGACCGTA AAGAAAAATA AGACAAGTT TTATCCGGCC  
 541 TTTATTCACA TTCTGCCCCG CCTGATGAAT GCTCATCCGG AATTCGTTAT GGCAATGAAA  
 601 GACGGTGAGC TGGTGTATG GGATAGTGTGTT CACCCCTGTT ACACCGTTT CCATGAGCAA  
 661 ACTGAAACGT TTTCATCGCT CTGGAGTGAA TACCAACGAGC ATTCGGCA GTTCTACAC  
 721 ATATATTGCG AAGATGTGGC GTGTTACCGT GAAAACCTGG CCTATTTCCC TAAAGGGTTT  
 781 ATTGAGAATA TGTTTTTCGT CTCAGCCAAT CCCTGGGTGA GTTTCACCAAG TTTTGATTITA  
 841 AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA CCATGGGCAA ATATTATACG  
 901 CAAGGCGACA AGGTGCTGAT GCGCGTGGCG ATTCAGGTTT ATCATGCCGT CTGTGATGGC  
 961 TTCCATGTCG GCAGAAATGCT TAATGAATTA CAACAGTACT GCGATGAGTG GCAGGGCGGG  
 1021 GCGTAAAGAT CTGGATCCGG CTTACTAAAAA GCCAGATAAC AGTATGCGTA TTGCGCGCT  
 1081 GATTTTGCG GTATAAGAT ATATACTGAT ATGTATAACCC GAAGTATGTC AAAAGAGGT  
 1141 GTGCTATGAA CGACGCTATT ACAGTGACAG TTGACAGCGA CAGCTATCAG TTGCTCAAGG  
 1201 CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC ATGCAGAAATG AAGCCCGTCG  
 1261 TCTGCGTGCC GAACGCTGG AAGCGGAAAA TCAGGAAGGG ATGGCTGAGG TCGCCCGGTT  
 1321 TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC TGGTAAATG CAGTTTAAGG  
 1381 TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT GGATGTACAG AGTGTATTA  
 1441 TTGACACGCC CGGGCGACGG ATGGTGATCC CCCTGGCCAG TGACGCTCTG CTGTCAGATA  
 1501 AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATCGGGGA TGAAAGCTGG CGCATGATGA  
 1561 CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGGA AGAAGTGGCT GATCTCAGCC  
 1621 ACCGGGAAAA TGACATCAA AACGCCATTA ACCTGATGTT CTGGGGAAATA TAAATGTCAG  
 1681 GCTCCCTTAT ACACAGGCCAG TCTGCAAGTC GACCATAGTG ACTGGATATG TTGTGTTTA  
 1741 CAGTATTATG TAGTCTGTTT TTTATGCAAAT ATCTAATTITA ATATATTGAT ATTTATATCA  
 1801 TTTTACGTTT CTCGTTTCAGC TTTCTTGTAC AAAGTGGTTG ATTCGAGGCT GCTAACAAAG  
 1861 CCCGAAAGGA AGCTGAGTTG GCTGCTGCCA CCGCTGAGCA ATAACTAGCA TAACCCCTTG  
 1921 GGGCCTCTAA ACGGGTCTTG AGGGGTTTTTG TGCTGAAAGG AGGAACATATA TCCGGATATC  
 1981 CACAGGACGG GTGTGGTCGC CATGATCGCG TAGTCGATAG TGGCTCCAAG TAGCGAAGCG  
 2041 AGCAGGACTG GGCGCGGCC AAAGCGTCG GACAGTGCCT CGAGAACGGG TCGCGATAGA  
 2101 AATTGCAATCA ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGGCGAT GCTGTCGGAA  
 2161 TGGACGATAT CCCGCAAGAG GCCCCGGAGT ACCGGCATAA CCAAGCCTAT GCCTACAGCA  
 2221 TCCAGGGTGA CGGTGCCAG GATGACGATG AGGGCATTGT TAGATTCAT ACACGGTGCC  
 2281 TGACTGCGTT AGCAATTAA CTGTAATAA CTACCGCATT AAAGCTTATC GATGATAAGC  
 2341 TGTCAAACAT GAGAATTCTT GAAGACGAAA GGGCCTCGTG ATACGCCAT TTTTATAGGT  
 2401 TAATGTCATG ATAATAATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG  
 2461 CGGAACCCCT ATTTGTTAT TTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA  
 2521 ATAACCTGA TAAATGCTTC AATAATATTG AAAAGGAAG AGTATGAGTA TTCAACATTT  
 2581 CCGTGTGCC CTTATTCCCT TTTTGCAGGC ATTTGCCTT CCTGTTTTG CTCACCCAGA  
 2641 AACGCTGGTG AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA-

FIGURE 37B

97/240

2701 ACTGGATCTC AACAGCGGTA AGATCCTGGA GAGTTTCGC CCCGAAGAAC GTTTCCAAT  
 2761 GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTGTG ACGCCGGCA  
 2821 AGAGCAACTC GGTCGCCGCA TACACTATT TCAGAATGAC TTGGTTGAGT ACTCACCACT  
 2881 CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC  
 2941 CATGAGTGT AACACTGCGG CCAACTTAAC TCTGACAACG ATCGGAGGAC CGAAGGAGCT  
 3001 AACCGCTTT TTGACACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAAACCGGA  
 3061 GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCCTGCAG CAATGGCAAC  
 3121 AACGTTGCGC AAACATTAA CTGGCGAACT ACTTACTCTA GCTTCCCAGG AACAAATTAAT  
 3181 AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTCTG CGCTCGGCCG TTCCGGCTGG  
 3241 CTGGTTTATT GCTGATAAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTG TCATTGCAAGC  
 3301 ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC  
 3361 AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCAATTG  
 3421 GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCAATTITA  
 3481 ATTTAAAAGG ATCTAGGTGA AGATCCTTT TGATAATCTC ATGACCAAAA TCCCTTAACG  
 3541 TGAGTTTTG TGCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA  
 3601 TCCTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAAGCGGT  
 3661 GGTTTGTGTTG CCGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAACGT GCTTCAGCAG  
 3721 AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA  
 3781 CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAAGTGG CTGCTGCCAG  
 3841 TGGCGATAAG TCGTGTCTT CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA  
 3901 GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC  
 3961 CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTCCCCG AAGGGAGAAA  
 4021 GGCAGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACCGA GGGAGCTTCC  
 4081 AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT GACTTGAGCG  
 4141 TCGATTTTG TGATGCTCGT CAGGGGGCG GAGCCTATGG AAAAACGCCA GCAACCGGC  
 4201 CTTTTTACGG TTCTGGCCT TTTGCTGGCC TTTTGTCACT ATGTTCTTTC CTGCGTTATC  
 4261 CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGTAGTGA GCTGATACCG CTCGCCAG  
 4321 CGAACGACCC GAGCGCAGCG AGTCAGTGA CGAGGAAGCG GAAGAGCGCC TGATGCGTA  
 4381 TTTTCTCCCTT ACGCATCTGT GCGGTATTC ACACCGCATA TATGGTGCACT TCTCACTACA  
 4441 ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA CGTGACTGGG  
 4501 TCATGGCTGC GCCCGACAC CCGCAACAC CCGCTGACGC GCCCTGACGG GCTTGTCTGC  
 4561 TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG GAGCTGCATG TGTCAGAGGT  
 4621 TTTCACCGTC ATCACCGAAA CGCGCAGGGC AGCTCGGGTA AAGCTCATCA GCGTGGCTGT  
 4681 GAAGCGATTAC ACAGATGTCT GCCTGTTCAT CCGCGTCCAG CTCGTTGAGT TTCTCCAGAA  
 4741 GCGTTAATGT CTGGCTCTG ATAAAGCGGG CCATGTTAAG GGCAGGTTTT TCCCTGTTGG  
 4801 TCACTGATGC CTCCGTAA GGGGGATTT TGTTCATGGG GGTAAATGATA CCGATGAAAC  
 4861 GAGAGAGGAT GTCACGATA CGGGTTACTG ATGATGAACA TGCCCGGTTA CTGGAACGTT  
 4921 GTGAGGGTAA ACAACTGGCG GTATGGATGC GGCAGGACCA GAGAAAATC ACTCAGGGTC  
 4981 AATGCCAGCG CTTGTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG CAGCATCCTG  
 5041 CGATGAGAT CCGGAACATA ATGGTGCAGG GCGCTGACTT CCGCGTTTCC AGACTTTACG  
 5101 AAACACGGAA ACCGAAGACCC ATTCACTGTT TTGCTCAGGT CGCAGACGTT TTGCAAGCAGC  
 5161 AGTCGCTTCA CGTTCGCTCG CGTATCGGT ATTCACTCTG CTAACCACTA AGGCAACCCCC  
 5221 GCCAGCCTAG CCGGGTCTC AACGACAGGA GCACGATCAT GCGCACCCCG GGCAGGACC  
 5281 CAACGCTGCC CGAGATGCGC CGCGTGCAGC TGCTGGAGAT GGCAGACGCG ATGGATATGT  
 5341 TCTGCCAAGG GTGGTTTGC GCATTACAG TTCTCCGAA GAATTGATTG GCTCCAATTG  
 5401 TTGGAGTGGT GAATCCGTTA GCGAGGTGCC GCGGCTTCC ATTCAAGGTC AGGTGGCCCG  
 5461 GCTCCATGCA CCGCGACGCA ACAGGGGAG GCAGACAAGG TATAGGGCGG CGCCTACAAT  
 5521 CCATGCCAAC CCGTTCCATG TGCTCGCCGA GGCAGCATAA ATCGCCGTGA CGATCAGCGG  
 5581 TCCAGTGATC GAAGTTAGGC TGGTAAGAGC CGCGAGCGAT CTTGAAAGCT GTCCCTGATG  
 5641 GTCGTCATCT ACCTGCCTGG ACAGCATGGC CTGCAACCGC GGCATCCCGA TGCCGCCGGA  
 5701 AGCGAGAAGA ATCATAATGG GGAAGGCCAT CCAGCCTCGC GTCGCGAACG CCAGCAAGAC  
 5761 GTAGCCCAGC GCGTCGGCCG CCATGCCGC GATAATGGCC TGCTTCTCGC CGAAACGTTT  
 5821 GGTGGCGGG ACGTGTGACGA AGGCTTGAGC GAGGGCGTGC AAGATTCCGA ATACCGCAAG  
 5881 CGACAGGCCG ATCATCGTCG CGCTCCAGCG AAAGCGGTCC TCGCCGAAAA TGACCCAGAG  
 5941 CGCTGCCGGC ACCTGTCCTA CGAGTTGCAT GATAAAGAAG ACAGTCATAA GTGCGCGAC  
 6001 GATAGTCATG CCCCAGCGCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC TCAAGGGCAT  
 6061 CGGTGATCG AGCCTCTCCC TTATGCAGT CCTGCATTAG GAAGCAGCCC AGTAGTAGGT  
 6121 TGAGGCCGTT GAGCACCGCC GCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA-

FIGURE 37C

98/240

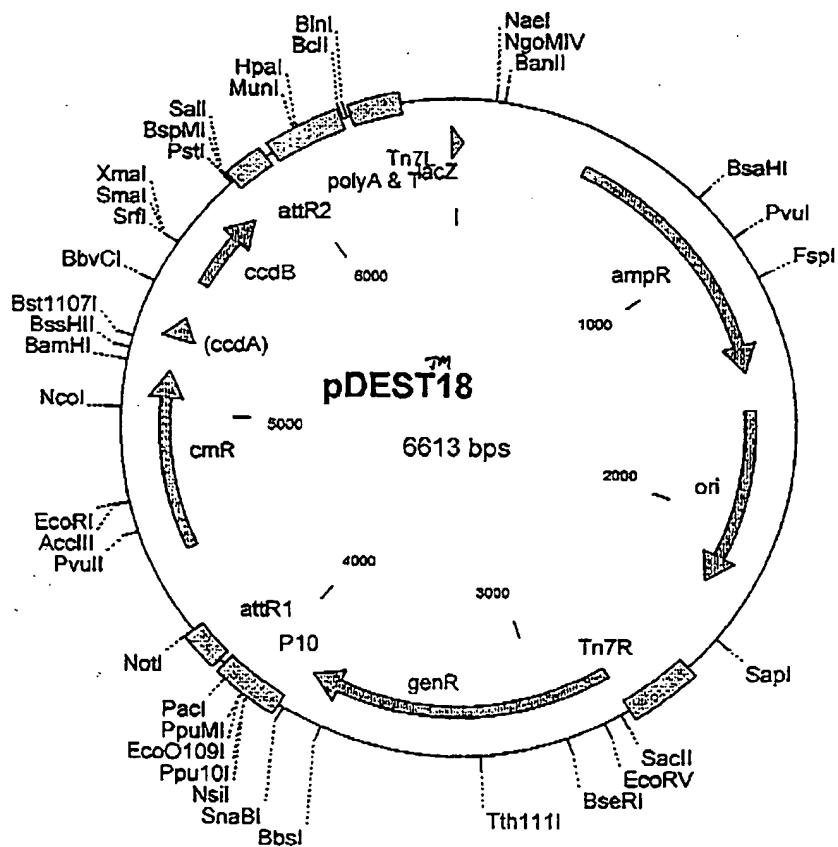
6181 GTCCCCCGGC CACGGGGCCT GCCACCATA CCACGCCGAA ACAAGCGCTC ATGAGCCGA  
6241 AGTGGCGAGC CCGATCTTCC CCATCGGTGA TGTGGCGAT ATAGGCGCCA GCAACCGCAC  
6301 CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCT

FIGURE 37D

**Figure 38A:** DESTIE

# **FastBac Transfer Vector with p10 Baculovirus Promoter**

1 gaagacctcg gccgtcgccgg cgcttgcggg tgggtctgac cccggatgaa gtggttcgea  
 cttctggac cggcagcgcc gcaacggcc accacgactg gggctactt caccaggcg  
 61 tcctcggtt tctggaaaggc gagcatcggt tggtcgccca ggactcttagc tatagttcta  
 aggagccaaa agacctccg ctcttagcaa acaagcggtt cctgagatcg atatcaagat  
 121 gtggttggct acgtatcgag caagaaaaata aaacgcccggc /cgcgttggag tcctgtgtgc  
 caccaccga tgcatacgcc ttcttttat ttgggggtt gcgcadccctc agaacaacgc  
 //  
 181 //tatttttaca aagatccaga aatcgccatc accttacaca aaaaaaaaaatatatgtt  
 //ataaaaatgt ttcttaatgtt ttatcgatgt tgaatgttgt tccccctgtt acctttaatcc  
 //  
 241 //cattttgagg atgcggggac cttsatcca acccaacaca atatattataa gtaaaatatgg mRNA  
 //ataaaaatcc tacggccctt qaaatataat tgggttgtgt tatataatat cattttatcc  
 //  
 301 //attttttat caaatcattt gtatattatataaataacta tactgtaaat tacattttat  
 //taataatataa gtttagtataa atatataatata attttatgtt atgacattt atgtaaaata  
 //  
 361 ttacaatgag gatcatcaca agtttgatca aaaaagctga acgagaaaacg taaaatgata  
 aatgttactc ctatgtt tcaaacatgt ttggggact tgcttttgc attttactat  
 //  
 Int → attR1



100 / 240

## pDEST18 6613 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
474..1449	ampR
1590..2244	ori
2738..3850	genR
4251..4127	attR1
4501..5160	CmR
5280..5364	inactivated ccdA
5502..5807	ccdB
5848..5972	attR2
6595..25	lacZ
1 GACGCGCCCT GTAGCGGCCG ATTAAGCGCG GC GGTTGTGG TGGTTACGCG CAGCGTGACC	
61 GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CT TTCTCGCC	
121 ACGTCGCCG GCTTTCCCCG TCAAGCTCTA AATC GGGGGC TCCCTTAGG GTTCCGATT	
181 AGTGCCTTAC GGCACCTCGA CCCAAAAAA CTTGATTAGG GTGATGGTT AC GTAGTGGG	
241 CCATCGCCCT GATAGACGGT TTTTCCCTT TTGACGTGG AGTCCACGTT CTTTAATAGT	
301 GGACTCTTGT TCCAAACTGG AACAAACACTC AACCTATCT CGGTCTATT CTTTGATTTA	
361 TAAGGGATTT TGCCGATTT GGCCTATTGG TTAAAAAATG AGCTGATT A CAAAAAATTT	
421 AACGCGAATT TTAACAAAAT ATTAACGTTT ACAATTTCAG GTGGCACTTT TC GGGGAAAT	
481 GTGCGCGAA CCCCTATTT TTTATTTTC TAAATACATT CAAATATGTA TCCGCTCATG	
541 AGACAATAAC CCTGATAAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA	
601 CATTCCGTG TCGCCCTTAT TCCCTTTT GCGCATT TT GCCTCCCTGT TTTGCTCAC	
661 CCAGAACGC TGGTGAAGT AAAAGATGCT GAAAGATCAGT TGGGTGCACG AGTGGGTTAC	
721 ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TCGCCCGA AGAACGTTT	
781 CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCC TATTGACGCC	
841 GGGCAAGAGC AACTCGGTG CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA	
901 CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCCTGCC	
961 ATAACCATGA GTGATAACAC TGCGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG	
1021 GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTA CTCGCCTTGA TCGTTGGAA	
1081 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG	
1141 GCAACAAACGT TGCGAAACT ATTAACGTT GAACTACTTA CTCTAGCTTC CGGGCAACAA	
1201 TTAATAGACT GGATGGAGGC GGATAAAAGTT GCAGGACAC TTCTGCGCTC GGCCCTTCCG	
1261 GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT	
1321 GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT	
1381 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG	
1441 CATTGGTAAC TGTCA GACCA AGTTTACTCA TATA TACTTT AGATTGATTT AAAACTTCAT	
1501 TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT	
1561 TAACGTGAGT TTTCGTTCCA CTGAGCGTC GACCCCGTAG AAAAGATCAA AGGATCTTCT	
1621 TGAGATCCTT TTTCGCG CGTAATCTGC TGCTTGAAA CAAAAAACC ACCGCTACCA	
1681 GCGGTGGTTT GTTGC GGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC	
1741 AGCAGAGCGC AGATACAAA TACTGCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC	
1801 AAGAACTCTG TAGCACGCC TACATACCTC GCTCTGCTAA TCCCTGTTACC AGTGGCTGCT	
1861 GCCAGTGGCG ATAAGTCGT TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG	
1921 GCGCAGCGGT CGGGCTGAAC GGGGGTTCG TGACACAGC CCAGCTTGA GCGAACGACC	
1981 TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG	
2041 AGAAAGGC GGACAGGTATCC GGTAAGCGGC AGGGTCGAA CAGGAGAGCG CACGAGGGAG	
2101 CTTCCAGGG GAAACGCCGT GTATCTTAT AGTCCTGTCG GTGGTGCCTA CCTCTGACTT	
2161 GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC	
2221 GCGGCCCTT TACGGTTCC TGGCTTTGC TGGCTTTTG CTCACATGTT CTTTCCTGCG	
2281 TTATCCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC	
2341 CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCTGATG	
2401 CGGTATTTTC TCCCTACGCA TCTGTGCGGT ATTTCACACC GCAGACCAGC CGCGTAACCT	
2461 GGCAAAATCG GTTACGGTTG AGTAATAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA-	

FIGURE 38B

101/240

2521 CAATAAAGTC TTAAACTGAA CAAAATAGAT CTAAACTATG ACAATAAAGT CTTAAACTAG  
 2581 ACAGAAATAGT TGTAACACTGA AATCAGTCCA GTTATGCTGT GAAAAGCAT ACTGGACTTT  
 2641 TGTTATGGCT AAAGCAAAC TTTCACTTTC TGAAGTGCAA ATTGCCCGTC STATTAAAGA  
 2701 GGGCGTGGC CAAGGGCATG GTAAAGACTA TATTGCGGGC GTTGTGACAA TTTACCGAAC  
 2761 AACTCCGCG CGGGGAAGCC GATCTCGCT GTAAACGAATT GTAGGTGGC GGTACTTGGG  
 2821 TCGATATCAA AGTGCATCAC TTCTTCCCGT ATGCCAACT TTGTATAGAG AGCCACTGCG  
 2881 GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG CACCAAGGCC GTTGGCCTCA  
 2941 TGCTTGAGGA GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGTG GCCGGAGACT  
 3001 GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA ACCTGGGCAG AACGTAAGCC  
 3061 GCGAGAGCGC CAACAACCGC TTCTTGGTCG AAGGCAGCAA GCGGATGAA TGTCTTACTA  
 3121 CGGAGCAAGT TCCCGAGGTA ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACSTCT  
 3181 CCGAACTCAC GACCGAAAAG ATCAAGAGCA GCCCCCATGG ATTTGACTTG GTCAGG3CCG  
 3241 AGCCTACATG TCGGAATGAT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCCACTG  
 3301 CCCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAACA  
 3361 TCGACCCACG GCGTAACGCG CTTGCTGCTT GGATGCCGA GGCATAGACT GTACAAAGAA  
 3421 ACAGTCATAA CAAGCCATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGCTCAA  
 3481 GGTTCTGGAC CAGTTGCGTG AGCGCATAACG CTACTTGCAT TACAGTTTAC GAACCGACA  
 3541 GGCTTATGTC AACTGGGTTG GTGCCCTCAT CCGTTTCCAC GGTGTGCGTC ACCCGCAAC  
 3601 CTTGGGCAGC AGCGAAGTCG AGGCATTCT GTCTGGCTG GCGAACGAGC GCAAGGTTTC  
 3661 GGTCTCCACG CATCGTCAGG CATTGGCGGC CTTGCTGTT TCCTACGGCA AGGTGCTGTG  
 3721 CACGGATCTG CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGCCTGGT  
 3781 GGTGCTGACC CGGGATGAAG TGGTCGCA CCTCGGTTTT CTGGAAGGCG AGCATCGTTT  
 3841 GTTCGCCCCAG GACTCTAGCT ATAGTTCTAG TGGTGGCTA CGTATCGAGC AAGAAAATAA  
 3901 AACGCCAAC CGCTTGGAGT CTTGTTGCT ATTTTACAA AGATTCAAGAA ATACGCATCA  
 3961 CTTACAACAA GGGGGACTAT GAAATTATGC ATTTGAGGA TGCCGGGACC TTTAATTCAA  
 4021 CCCAACACAA TATATTATAG TTAAATAAGA ATTATTTATC AAATCATTG TATATTAAATT  
 4081 AAAATACTAT ACTGTAAATT ACATTTATT TACAATGAGG ATCATCACAA GTTGTACAA  
 4141 AAAAGCTGAA CGAGAAACGT AAAATGATAT AAATATCAAT ATATTAATT AGATTTGCA  
 4201 TAAAAAACAG ACTACATAAT ACTGTAAAAC ACAACATATC CAGTCACTAT GGCGCCCGCT  
 4261 AAGTGGCAG CATCACCCGA CGCACTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC  
 4321 TTCGCAGAAT AAATAAAATCC TGGTGTCCCT GTTGATACCG GGAAGCCCTG GGCCAACTTT  
 4381 TGGCGAAAAT GAGACGTTGA TCGGCACGTA AGAGGTTCCA ACTTTCACCA TAATGAAATA  
 4441 AGATCACTAC CGGGCGTATT TTTGAGTTA TCGAGATTTT CAGGAGCTAA GGAAGCTAAA  
 4501 ATGGAGAAAA AAATCACTGG ATATACCACC GTTGATATAT CCCATGGCA TCGTAAGGAA  
 4561 CATTGGAGG CATTTCAGTC AGTTGCTCAA TGTACCTATA ACCAGACCGT TCAGCTGGAT  
 4621 ATTACGGCT TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTTATCC GGCCTTATT  
 4681 CACATTCTTG CCCGCCTGAT GAATGCTCAT CCGGAATTCC GTATGGCAAT GAAAGACGGT  
 4741 GAGCTGGTGA TATGGGATAG TGGTACCCCT TGGTACACCG TTTTCCATGA GCAAACGTGAA  
 4801 ACGTTTCAT CGCTCTGGAG TGAATACCAC GACGATTCC GGCAGTTCT ACACATATAT  
 4861 TCGCAAGATG TGGCGTGTGA CGGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTATTGAG  
 4921 AATATGTTT TCGTCTCAGC CAATCCCTGG GTGAGTTCA CCAGTTTGA TTTAAACGTG  
 4981 GCCAAATATGG ACAACTTCTT CGCCCCCGTT TTCACCATGG GCAAATATTA TACGCAAGGC  
 5041 GACAAGGTGC TGATGCCGCT GGCAGATTCA GTTCATCATG CCGTCTGTGA TGGCTTCCAT  
 5101 GTCGGCAGAA TGCTTAATGA ATTACAACAG TACTGCGATG AGTGGCAGGG CGGGGCCTAA  
 5161 ACGCGTGGAT CGGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTTGCG CGCTGATTTT  
 5221 TCGGTATAA GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAG AGGTGTGCTA  
 5281 TGAAGCAGCG TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCATATA  
 5341 TGATGTCAT ATCTCCGGTC TGGTAAGCAG ACCATGCG AATGAAGCCC GTCGTC-GCG  
 5401 TGCGGAACGC TGGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTGCGCC GGTTTATTGA  
 5461 AATGAACGGC TCTTTGCTG ACGAGAACAG GGACTGGTGA AATGCAGTT AAGGTTTACA  
 5521 CCTATAAAAG AGAGAGCCGT TATCGTCTGT TTGTTGGATGT ACAGAGTGT ATTATTCACA  
 5581 CGCCCGGGCG ACGGATGGTG ATCCCCCTGG CCAGTCACG TCTGCTGTCA GATAAAAGTCT  
 5641 CCCGTGAAC TTAACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCCACCG  
 5701 ATATGGCCAG TGTGCCGGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCCCG  
 5761 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTGCTGGGG AATATAAATG TCAGGCTCCC  
 5821 TTATACACAG CCAGTCCTGCA GGTCGACCAT AGTGACTGGA TATGTTGTT TTTACASTAT  
 5881 TATGAGTCT GTTTTTATG CAAAATCTAA TTTAATATAT TGATATTAT ATCATTTTAC  
 5941 GTTCTCGTT CAGTTTCTT GTACAAAGTG GTGATAGCTT GTCGAGAAGT ACTAGAAGT-

FIGURE 38C

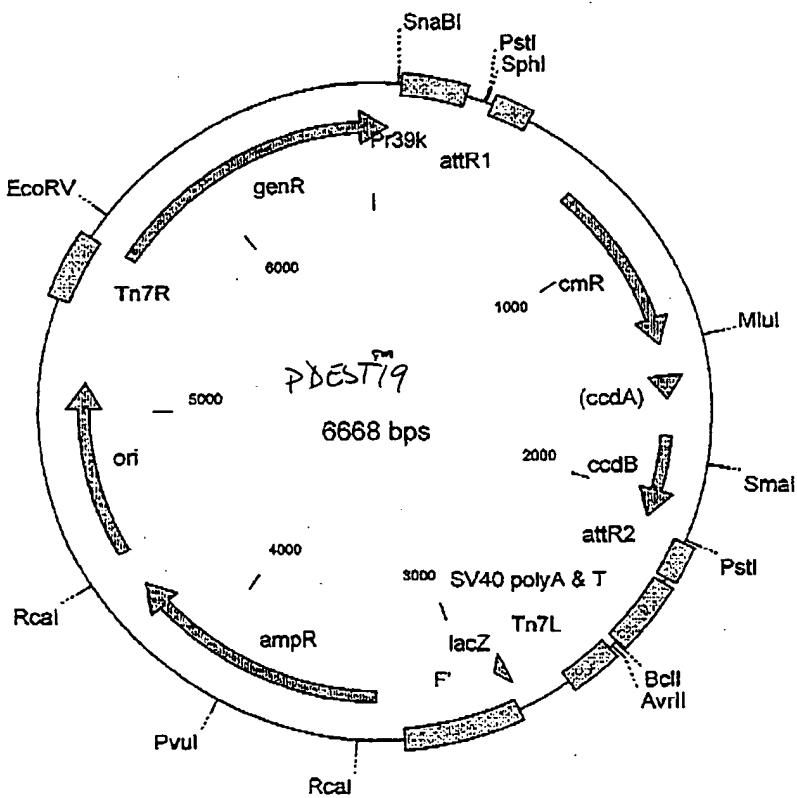
102/260

6001 CATAATCAGC CATACCACAT TTGTAGAGGT TTTACTTGCT TTAAAAAAACC TCCCACACCT  
6061 CCCCTGAAAC CTGAAACATA AAATGAATGC AATTGTTGGTT GTTAACCTTGT TTATTGCAGC  
6121 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTC ACAAAATAAAG CATTTC  
6181 ACTGCATTCT AGTTGTGGTT TGTCCTAACT CATCAATGTA TCTTATCATG TCTGGATCTG  
6241 ATCACTGCTT GAGCCTAGGA GATCCGAACC AGATAAGTGA AATCTAGTTC CAAACTATTT  
6301 TGTCACTTTT AATTTTCGTA TTAGCTTAGC ACGCTACACC CAGTTCCCAT CTATTTGTC  
6361 ACTCTTCCCT AAATAATCCT TAAAAACTCC ATTTCCACCC CTCCCAGTTC CCAACTATTT  
6421 TGTCCGCCCA CAGCGGGGCA TTTTCTTCC TGTTATGTTT TTAATCAAAC ATCCTGCCAA  
6481 CTCCATGTGA CAAACCGTCA TCTTCGGCTA CTTTTCTCT GTCACAGAAT GAAAATTTT  
6541 CTGTCATCTC TTCGTTATTA ATGTTTGTAA TTGACTGAAT ATCAACGCTT ATTTGCAGCC  
6601 TGAATGGCGA ATG

FIGURE 38D

103/240

1 ggtgacggcc tcatcttccc attgttaacgt aaatggcaac ttgttagatga acgcgcgtgtc  
ccactgcggc agtagaaaagg taacattgca tttaccgttg aacatctact tgcgcgacag  
  
61 aaaaaaacgg ccagtttctt ccacaaaactc ggcgcacggct gtctcgtaaa cttttgcgtc  
tttttggcc ggtcaaagaa ggtgttttag cgcgccgca cagagcattt gaaaacgcag  
  
121 // gcaacaatcg cgatgacetc gggttatgga aaaaaaaaaaattttttctt aaaaaagtgt cgttcatgtc  
// cgttgttagc qctactggag caccataacct taaaaaaaaga ttttttcaca gcaagtacag  
  
181 // ggcggccggcg ttgcgcgtcc ggtacgcgcg acgggcacac agcaggacag ctttgcgtccgg  
// cccggccgc aagcgcgagg ccatgcgcg tgccccgtgtc tggttttcgc ggaacaggcc  
// **ATP**  
  
241 ctgcattatc ataaaacaatc ctgcaggcat gcaagctgga tcatccaaag ttgttacaaa  
gagctaatacg tattttgttag gacgtccgtc cgttcgaccc agtagtgttc aaacatgttt  
Int V



104/240

## pDEST19 6668 bp (rotated to position 1000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
515..391	attR1
765..1424	CmR
1544..1628	inactivated ccdA
1766..2071	ccdB
2112..2236	attR2
2852..2895	lacZ
3344..4319	ampR
4460..5114	ori
5608..52	genR

1 AGTGGTTCGC ATCCTCGGTT TTCTGGAAGG CGAGCATCGT TTGTCGCC AGGACTCTAG  
 61 CTATAGTTCT AGTGGTTGGC TACGTATATC AAATACTTGT AGGTGACGCC GTCATCTTC  
 121 CATTGTAACG TAAATGGCAA CTTGTAGATG AACCGCCTGT CAAAAAACCG GCCAGTTCT  
 181 TCCACAAACT CGCGCACGGC TGTCTCGTAA ACTTTTGCCT CGCAACAATC GCGATGACCT  
 241 CGTGGTATGG AAATTTTTTC TAAAAAAAGTG TCGTTCATGT CGGCGCGGG CGCGTTCGCG  
 301 CTCCGGTACG CGCGACGGGC ACACAGCAGG ACAGCCTTGT CGGCTCGAT TATCATAAAC  
 361 AATCCTGCAG GCATGCAAGC TCGGATCATC ACAAGTTTGT ACAAAAAAAGC TGAACGAGAA  
 421 ACGTAAAATG ATATAAATAT CAATATATTA AATTAGATTT TGCAAAAAA ACAGACTACA  
 481 TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGGC CGCTAAGTTG GCAGCATCAC  
 541 CCGACGCACT TTGCGCCGAA TAAATACCTG TGACGGAAGA TCACTTCGCA GAATAAATAA  
 601 ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC CCTGGGCCAA CTTTTGGCGA AAATGAGACG  
 661 TTGATCGGCA CGTAAGAGGT TCCAACCTTC ACCATAATGA AATAAGATCA CTACCGGGCG  
 721 TATTTTTTGA GTTATCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG AAAAAAAATCA  
 781 CTGGATATAC CACCGTTGAT ATATCCAAT GGATCGTAA AGAACATTG GAGGCATTTC  
 841 AGTCAGTTGC TCAATGTACC TATAACCAGA CGGTTCACT GGATATTACG GCCTTTTAA  
 901 AGACCGTAAA GAAAAATAAG CACAAGTTT ATCCGGCCTT TATTACATT CTTGCCGCC  
 961 TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA CGGTGAGCTG GTGATATGGG  
 1021 ATAGTGTCA CCCTTGTAC ACCGTTTCC ATGAGCAAAC TGAAACGTT TCATCGCTCT  
 1081 GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT ATATTGCAA GATGTGGCGT  
 1141 GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTAT TGAGAATATG TTTTCTGCT  
 1201 CAGCCAATCC CTGGGTGAGT TTCACCACTT TTGATTAAA CGTGGCAAT ATGGACAAC  
 1261 TCTTCGCCCC CGTTTCACC ATGGGCAAAT ATTATACGCA AGGCGACAAG GTGCTGATGC  
 1321 CGCTGGCGAT TCAGGTTCAT CATGGCTCT GTGATGGCTT CCATGTCGGC AGAATGCTTA  
 1381 ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC GTAAACGCGT GGATCCGGCT  
 1441 TACTAAAAGC CAGATAAACAG TATGCGTATT TGCAGCCTGA TTTTTGCGGT ATAAGAATAT  
 1501 ATACTGATAT GTATACCCGA AGTATGCAA AAAGAGGTTG GCTATGAAAGC AGCGTATTAC  
 1561 AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA TATATGATGT CAATATCTCC  
 1621 GGTCTGGTAA GCACAACCAT GCAGAATGAA GCCCCTCGTC TGCGTGCAGA ACGCTGGAAA  
 1681 GCGGAAAATC AGGAAGGGAT GGCTGAGGTC GCCCGGTTTA TTGAAATGAA CGGCTTTT  
 1741 GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT TACACCTATA AAAGAGAGAG  
 1801 CCGTTATCGT CTGTTGTCG ATGTACAGAG TGATATTATT GACACGCCGG GGCACGGAT  
 1861 GGTGATCCCC CTGGCCAGTG CACGTCGCT GTCACTGAA GTCTCCCGTG AACTTTACCC  
 1921 GGTGGTGCAT ATCGGGATG AAAGCTGGC CATGATGACC ACCGATATGG CCAGTGTGCC  
 1981 GGTCTCCGTT ATCGGGGAAAG AAGTGGCTGA TCTCAGGCCAC CGCGAAAATG ACATAAAAAA  
 2041 CGCCATTAAC CTGATGTTCT GGGGAATATA AATGTCAGGC TCCCTTATAC ACAGCCAGTC  
 2101 TGCAGGTCGA CCTAATGAC TGGATATGTT GTGTTTTACA GTATTATGTA GTCTGTTTT  
 2161 TATGCAAAAT CTAAATTAAAT ATATTGATAT TTATATCATT TTACGTTCT CGTTCAGCTT  
 2221 TCTTGTACAA AGTGGTGAATC GAGAAGTACT AGAGGATCAT AATCAGCCAT ACCACATTG  
 2281 TAGAGTTTT ACTTGCTTTA AAAAACCTCC CACACCTCCC CCTGAACCTG AAACATAAAA  
 2341 TGAATGCAAT TGTTGTTGTT AACTTGTTTA TTGAGCTTA TAATGGTTAC AAATAAAGCA  
 2401 ATAGCATCAC AAATTTCAAA AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTGT  
 2461 CCAAACTCAT CAATGTATCT TATCATGTCT GGATCTGATC ACTGCTTGAG CCTAGGAGAT  
 2521 CCGAACCGAGA TAAGTGAAT CTAGTCCAA ACTTTTGT CATTTTAAT TTTCGTATTAA  
 2581 GCTTACGACG CTACACCCAG TTCCCATCTA TTTTGTCACT CTTCCCTAAA TAATCCTTAA-

FIGURE 39B

105/240

2641 AAACCTCCATT TCCACCCCTC CCAGTTCCA ACTATTTGT CCGCCCACAG CGGGGCATT  
 2701 TTCTTCTGT TATGTTTTA ATCAAACATC CTGCCAACTC CATGTGACAA ACCGTCACT  
 2761 TCGGCTACTT TTTCTCTGTC ACAGAATGAA AATTTTCTG TCATCTCTTC GTTATTAATG  
 2821 TTGTAATTG ACTGAATATC AACGCTTATT TGCAGCCTGA ATGGCGAATG GACGCGCCCT  
 2881 GTAGCGGCAG ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTG  
 2941 CCAGCCCCCT AGCGCCCGCT CCTTCCGCTT TCTTCCCTTC CTTTCTGCC ACGTTCGCCG  
 3001 GCTTCCCCG TCAAGCTTA AATCGGGGGC TCCCTTAGG GTTCCGATTG AGTGTCTTAC  
 3061 GGCACCTCGA CCCCAAAAAA CTTGATTAGG GTGATGGTC AGCTAGTGGG CCATGCCCT  
 3121 GATAGACGGT TTTTCGCCCC TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT  
 3181 TCCAAACTGG AACAAACACTC AACCTATCT CGGTCTATTG TTTTGTATTA TAAGGGATT  
 3241 TGCCGATTTG GGCCTATTGG TTAAAAAATG AGCTGATTG AAAAAAATT AACCGGAATT  
 3301 TTAACAAAAT ATTAACGTTT ACAATTTCAG GTGGCACTTT CGGGGAAAT GTGCGCGAA  
 3361 CCCCTATTG TTTATTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC  
 3421 CCTGATAAAAT GCTTCATAAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTCCGTG  
 3481 TCGCCCTTAT TCCCTTTTT GCGGCATTG GCCTTCTGT TTTTGTCTAC CCAGAACGC  
 3541 TGGTAAAGT AAAAGATGCT GAAGATCAGT TGGTGCACG AGTGGGTTAC ATCGAATCTG  
 3601 ATCTAACAG CGGTAAGATC CTTGAGAGTT TTGCCCCGA AGAACGTTT CCAATGATGA  
 3661 GCACCTTAA AGTTCTGCTA TGTGGCGCG TATTATCCCG TATTGACGCC GGGCAAGAGC  
 3721 AACTCGGTCG CCGCATACAC TAITCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG  
 3781 AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGTGCC ATAACCATGA  
 3841 GTGATAACAC TGCGGCCAAC TTACTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG  
 3901 CTTTTTGCA CAACATGGGG GATCATGTA CTCGCCTGTA TCGTTGGGAA CCGGAGCTGA  
 3961 ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAAACGT  
 4021 TGCGAAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC CGGCAACAA TTAATAGACT  
 4081 GGATGGAGGC GGATAAAAGTT GCAGGACAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT  
 4141 TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG  
 4201 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA  
 4261 TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCTC ACTGATTAAG CATTGGTAAC  
 4321 TGTCAAGACCA AGTTTACTCA TATATACTTT AGATTGATTG AAAACTTCAT TTTTAATT  
 4381 AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT  
 4441 TTTCGTTCCA CTGAGCTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT  
 4501 TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTT  
 4561 GTTTGCCGGA TCAAGACTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC  
 4621 AGATAACAAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG  
 4681 TAGCACCGCC TACATACCTC GCTCTGCTAA TCCCTGTTACC AGTGGCTGCT GCCAGTGGCG  
 4741 ATAAGTCGTG TCTTACGGG TTGGACTCAA GAGGATAGTT ACCGATAAG GCGCAGCGGT  
 4801 CGGGCTGAAC GGGGGGTTCG TGACACAGC CCAGCTTGGA GCGAACGAC TACACCGAAC  
 4861 TGAGATAACCT ACAGCTGAG CATTGAGAAA GCGCCACGCT TCCGAAGGG AGAAAGGCGG  
 4921 ACAGGTATCC GTAAAGCGGC AGGGTCGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG  
 4981 GAAACGCCCTG GTATCTTAT AGTCCTGTCG GTTTCGCCA CCTCTGACTT GAGCGTCGAT  
 5041 TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTT  
 5101 TACGGTTCCCT GCCCTTTG TGGCTTTG CTCACATGTT CTTTCTGCC TTATCCCCCTG  
 5161 ATTCTGTGGA TAACCGTATT ACCGCCTTGT AGTGAGCTGA TACCGCTCGC CGCAGCCGAA  
 5221 CGACCGAGCG CAGCGAGTC GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC  
 5281 TCCCTACGCA TCTGTGCGGT ATTTACACCC GCAGACCGAGC CGCGTAACCT GGCAAAATCG  
 5341 GTTACGGTTG AGTAATAAAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA CAATAAGTC  
 5401 TAAACTGAA CAAAATAGAT CTAAACTATG ACAATAAAAGT CTTAAACTAG ACAGAATAGT  
 5461 TGTAAACTGA AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TGTTATGGCT  
 5521 AAAGCAAACCT TTTCATTTTC TGAAGTGAA ATTGCCCCGTC GTATTAAAGA GGGCGTGGC  
 5581 CAAGGGCATG GTAAAGACTA TATTGCGGC GTTGTGACAA TTTACCGAAC AACTCCGCGG  
 5641 CGGGGAAGCC GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA  
 5701 AGTCATCAC TTCTTCCCGT ATGCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC  
 5761 CGTAATCTGC TTGACGCTAG ATCACATAAG CACCAAGCGC GTTGGCCTCA TGCTTGAGGA  
 5821 GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCCGGAGACT GCGAGATCAT  
 5881 AGATATAGAT CTCACTACGC GGCTGCTAA ACCTGGGCAG AACGTAAGCC GCGAGAGCGC  
 5941 CAACAACCGC TTCTTGGTC AAGGCAGCAA GCGCGATGAA TGTCTTAACCA CGGAGCAAGT  
 6001 TCCCGAGGTA ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC  
 6061 GACCGAAAAG ATCAAGAGCA GCGCCGATGG ATTTGACTTG GTCAGGGCG AGCCTACATG-

FIGURE 39C

106/240

6121 TGCAGATGAT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCG  
6181 TAACATCGTT GCTGCTGCGT AACATCGTT CTGCTCCATA ACATCAAACA TCGACCCACG  
6241 GCGTAACCGCG CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA  
6301 CAAGCCATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC  
6361 CAGTTGCGTG AGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAACAA GGCTTATGTC  
6421 AACTGGGTTC GTGCCCTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC  
6481 AGCGAAGTCG AGGCATTCT GTCCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG  
6541 CATCGTCAGG CATTGGCGGC CTTGCTGTTT TTCTACGGCA AGGTGCTGTG CACGGATCTG  
6601 CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGCCGGT GGTGCTGACC  
6661 CCGGATGA

FIGURE 39D

107/240

Figure 40A: pDEST20 Glutathione-S-transferase Fusion with Polyhedron Promoter for Baculovirus Expression

430 ggc tac gta tac tcc gga ata tta ata gat cat gga gat aat taa aat gat  
 ccg atg cat atg agg cct tat aat tat cta gta cct cta tta att tta cta //

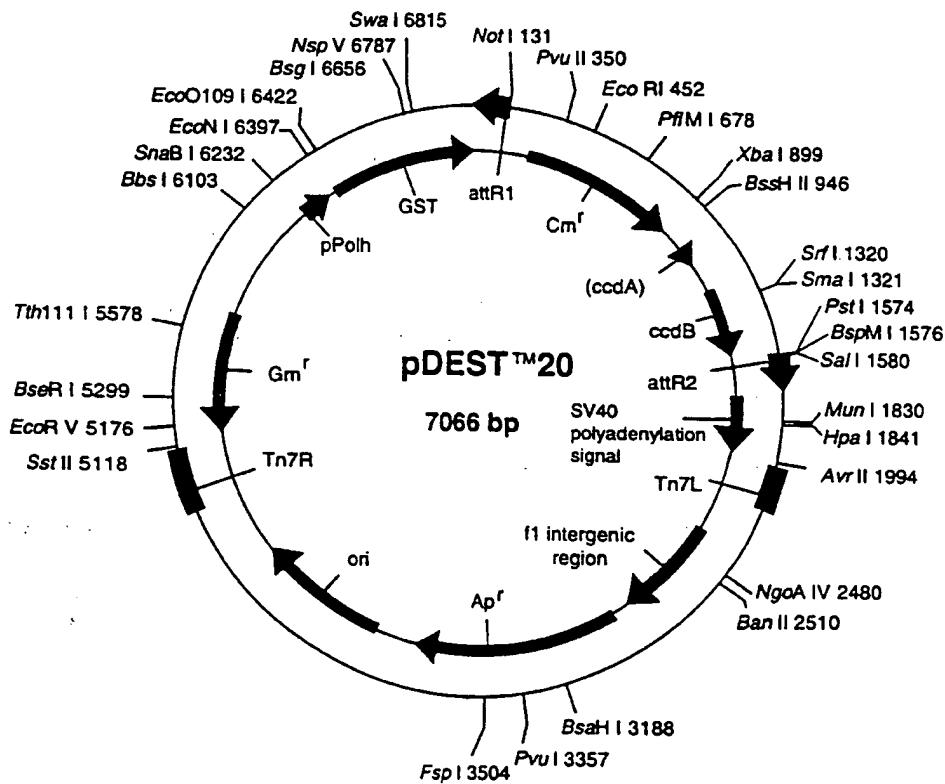
481 aac cat ctc gca aat aaa taa gta ttt tac tgt ttt cgt aac agt ttt gta  
 //tgc gta gag cgt tta ttt att cat aaa atg aca aaa gca ttg tca aaa cat //

532 ata aaa aaa cct ata aat att ccg gat tat tca tac cgt ccc acc atc ggg  
 //tat ttt ttt gga tat tta taa ggc cta ata agt atg gca ggg tgg tag ccc

Start Transl. M → A P T - - - GST - -  
 583 cgc gga tcc atg gct cct ata cta ggt tat tgg aaa att aag ggc ctt gtg —//  
 gcg cct agg tac cgg gga tat gat cca ata acc ttt taa ttc cgg gaa cac

1246 S D L V P R H N Q T S L Y K K A  
 //tcc gat ctg gtt ccg cgt cat aat caa aca agt ttg tac aaa aaa gct gaa  
 agc cta gac caa ggc gca gta tta gtt tgt tca aac atg ttt ttt cga cct

1297 cga gaa acg taa aat gat ata aat atc aat ata tta aat tag at  
 gct ctt tgc att tta cta tat tta tag tta tat aat tta atc ta



108/240

## pDEST20 7066 bp (rotated to position 5800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
592..1263	GST
1397..1273	attR1
1506..2165	CmR
2285..2369	inactivated ccdA
2507..2812	ccdB
2853..2977	attR2
4214..5064	ampR
5263..5843	ori

1 CCACTGCGCC GTTACCACCG CTGCGTCGG TCAAGGTTCT GGACCAGTTG CGTGAGCGCA  
 61 TACGCTACTT GCATTACAGT TTACGAACCG AACAGGCTTA TGTCAACTGG GTTCGTGCCT  
 121 TCATCCGTTT CCACGGTGTG CGTCACCCGG CAACCTTGGG CAGCAGCGAA GTCGAGGCAT  
 181 TTCTGTCCTG GCTGGCGAAC GAGCGCAAGG TTTCGGTCTC CACGCATCGT CAGGCATTGG  
 241 CGGCCCTGCT GTTCTTCTAC GGCAAGGTGC TGTGCACGGA TCTGCCCTGG CTTCAGGAGA  
 301 TCGGAAGACC TCGGCCGTG CGGCCTTGC CGGTGGTGT GACCCCGGAT GAAGTGGTTC  
 361 GCATCCTCGG TTTTCTGGAA GGCAGCATC GTTGTTCGC CCAGGACTCT AGCTATAGTT  
 421 CTAGTGGTTG GCTACGTATA CTCCGGAATA TTAATAGATC ATGGAGATAA TTAAAATGAT  
 481 AACCATCTCG CAAATAAATA AGTATTTAC TGTTTCGTA ACAGTTTGT AATAAAAAAA  
 541 CCTATAAATA TTCCGGATTA TTCATACCGT CCCACCATCG GGCGCGGATC CATGGCCCCCT  
 601 ATACTAGGTT ATTGGAAAAT TAAGGGCTT GTGCAACCCA CTCGACTTCT TTTGGAATAT  
 661 CTTGAAGAAA AATATGAAGA GCATTGTAT GAGCGCGATG AAGGTGATAA ATGGCGAAAC  
 721 AAAAGTTTG AATTGGGTTT GGAGTTCCC AATCTCCTT ATTATATTGA TGGTGTATGTT  
 781 AAATTAACAC AGTCTATGGC CATCATACTG TATATAGCTG ACAAGCACAA CATGTTGGGT  
 841 GGTTGTCCTA AAGAGCGTGC AGAGATTCA ATGCTTGAAG GAGCGGTTTT GGATATTAGA  
 901 TACGGTGTTC CGAGAATTGC ATATAGTAAA GACTTTGAAA CTCTCAAAGT TGATTTCTT  
 961 AGCAAGCTAC CTGAAATGCT GAAAATGTTG GAAGATCGTT TATGTCATAA AACATATTAA  
 1021 AATGGTGATC ATGTAACCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA TGTTGTTTA  
 1081 TACATGGACC CAATGTGCCT GGATGCGTTG CCAAAATTAG TTTGTTTAA AAAACGTATT  
 1141 GAAGCTATCC CACAAATTGA TAACTACTTG AAATCCAGCA AGTATATAGC ATGGCCTTTG  
 1201 CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCTC CAAAATCGGA TCTGGTTCCG  
 1261 CGTCATAATC AAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT  
 1321 ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA CATAACTTG TAAAACACAA  
 1381 CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG  
 1441 GCTCGTATGT TGTGTGGATT TTGAGTTAGG ATCCGGCGAG ATTTTCAGGA GCTAAGGAAG  
 1501 CTAAAATGGA GAAAAAAATC ACTGGATATA CCACCGTTGA TATATCCCAA TGGCATCGTA  
 1561 AAGAACATTT TGAGGCATTG CAGTCAGTTG CTCATGTAC CTATAACCA ACCGTTTCAGC  
 1621 TGGATATTAC GCCCTTTTA AAGACCGTAA AGAAAAAATAA GCACAGTTT TATCCGGCCT  
 1681 TTATTACAT TCTTGCCCGC CTGATGAATG CTCATCCGGA ATTCGTTATG GCAATGAAAG  
 1741 ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGTGA CACCGTTTC CATGAGCAAA  
 1801 CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG TTTCTACACA  
 1861 TATATTGCA AGATGTGGCG TGTACCGGT AAAACCTGGC CTATTTCCCT AAAGGGTTTA  
 1921 TTGAGAATAT GTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTCACCAGT TTGATTTAA  
 1981 ACGTGGCCAA TATGGACAAAC TTCTTCCGCC CGGTTTTCAC CATGGGAAA TATTATACGC  
 2041 AAGGCACAA GGTGCTGATG CCGCTGGCGA TTCAAGTTCA TCATGCCGTC TGTGATGGCT  
 2101 TCCATGTCGG CAGAATGCTT AATGAATTAC AACAGTACTG CGATGAGTGG CAGGGCGGGG  
 2161 CGTAACTTAG AGGATCCGGC TTACTAAAAG CCAGATAACA GTATGCGTAT TTGCGCGCTG  
 2221 ATTTTGCGG TATAAGAATA TATACTGATA TGTATACCG AAGTATGTCA AAAAGAGGTG  
 2281 TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC AGCTATCAGT TGCTCAAGGC  
 2341 ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAATGA AGCCCGTCGT  
 2401 CTGGCTGCCG AACGCTGGAA AGCGGAAAAT CAGGAAGGG TGGCTGAGGT CGCCCGGTTT  
 2461 ATTGAAATGA ACGGCTCTTT TGCTGACGAG AACAGGGACT GGTGAAATGC AGTTAAGGT  
 2521 TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTGTG GATGTACAGA GTGATATTAT  
 2581 TGACACGCCG GGGCGACGGA TGGTGTACCC CCTGGCCAGT GCACGCTCTG TGTCAGATAA  
 2641 AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCGGGAT GAAAGCTGGC GCATGATGAC-

Figure 40B

109/240

2701 CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGAA GAAAGTGGCTG ATCTCAGCCA  
 2761 CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTT TC GGGAAATAT AAATGTCAGG  
 2821 CTCCCTTATA CACAGCCAGT CTGCAGGTCG ACCATAGTGA CTGGATATGT TGTGTTTAC  
 2881 AGTATTATGT AGTCTGTTT TTATGCAAAA TCTAATTAA TATATTGATA TTTATATCAT  
 2941 TTTACGTTTC TCGTTCAGCT TTCTTGACA AAGTGGTTG ATAGCTTGT GAGAAGTACT  
 3001 AGAGGATCAT ATCAGCCAT ACCACATTG TAGAGGTTT ACTTGTCTTA AAAAACCTCC  
 3061 CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTTGTTGTT AACTGTTTA  
 3121 TTGAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA AATAAAGCAT  
 3181 TTTTTCACT GCATTCTAGT TGTGGTTGT CCAAACCTCAT CAATGTATCT TATCATGTCT  
 3241 GGATCTGATC ACTGCTTGAG CCTAGGAGAT CCGAACAGA TAAAGTAAAT CTAGTTCCAA  
 3301 ACTATTTGT CATTTTTAAT TTTCTGATTA GCTTACGACG CTACACCCAG TTCCCATCTA  
 3361 TTTTGTCACT CTTCCCTAAA TAATCCCTAA AAACCTCCATT TCCACCCCTC CCAGTTCCCA  
 3421 ACTATTTGT CGGCCACAG CGGGGATTT TTCTTCCCTGT TATGTTTTA ATCAAACATC  
 3481 CTGCCAACTC CATGTGCAA ACCGTGATCT TCGGCTACTT TTTCTGTC ACAGAATGAA  
 3541 AATTTTCTG TCACTCTTC GTTATTAATG TTGTAATTG ACTGAATATC AACGTTTATT  
 3601 TGCAGCCTGA ATGGCAATG GACGCCCT GTAGCGCGC ATTAAGCGCG GCGGGTGTGG  
 3661 TGGTACGCG CAGCGTACCC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTCGCTT  
 3721 TCTTCCCTTC CTTTCTGCC ACGTTCGCC GCTTCCCCG TCAAGCTCTA AATCGGGGGC  
 3781 TCCCTTAGG GTTCCGATT AGTCTTAC GGCACCTCGA CCCCCAAAAA CTTGATTAGG  
 3841 GTGATGGTTC ACGTAGTGGG CCATGCCCT GATAGACGGT TTTTCGCCCT TTGACGTTGG  
 3901 AGTCCACGTT CTITAATAGT GGACTCTTGT TCCAAACTGG AACAAACACTC AACCTATCT  
 3961 CGGTCTATTTC TTTTGATTTA TAAGGGATT TGCGATTTT GGCCTATTGG TTAAAAAATG  
 4021 AGCTGATTTA AAAAAAATT AACCGGAATT TTAACAAAAT ATTAACGTTT ACAATTTCAG  
 4081 GTGGCACTTT TCAGGGAAAT GTGCGCGGAA CCCCTATTG TTTATTTTC TAAATACATT  
 4141 CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAAT GCTTCAATAA TATTGAAAAA  
 4201 GGAAGAGTAT GAGTATTCAA CATTCCGTG TCGCCCTTAT TCCCTTTTT GCGGCATTTC  
 4261 GCCTTCTGT TTTTGCTCAC CCAGAAACGC TGGTAAAGT AAAAGATGCT GAAGATCAGT  
 4321 TGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAAGATC CTTGAGAGTT  
 4381 TTGGCCCCGA AGAACGTTT CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGCG  
 4441 TATTATCCCCG TATTGACGCC GGGCAAGAGC AACTCGGTG CCGCATACAC TATTCTCAGA  
 4501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA  
 4561 GAGAATTATG CAGTGTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA  
 4621 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTAA  
 4681 CTGGCTTGA TCGTTGGAA CGGGAGCTGA ATGAAGCCAT ACCAACGAC GAGCGTGACA  
 4741 CCACGATGCC TGTAGCAATG GCAACAAACGT TCGCACAACCT ATTAACIGGC GAACTACTTA  
 4801 CTCTAGCTTC CGGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAAGTT GCAGGACCAC  
 4861 TTCTGCGCTC GGCCCTTCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC  
 4921 GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG  
 4981 TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA  
 5041 TAGGTGCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT  
 5101 AGATTGATT AAAACTTCAT TTTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTTGATA  
 5161 ATCTCATGAC CAAAATCCCT TAACGTGAGT TTCTGTTCCA CTGAGCGTCA GACCCCGTAG  
 5221 AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA  
 5281 CAAAAAAACC ACCGCTACCA CGGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT  
 5341 TTCCGAGGTT AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC  
 5401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA  
 5461 TCCGTCTTACG AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA  
 5521 GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGTTCG TGACACACAGC  
 5581 CCAGCTTGGG GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA  
 5641 GCGCCACGCT TCCCGAAGGG AGAAAGCGG ACAGGTATCC GGTAAAGCGGC AGGGTCGGAA  
 5701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTAT AGTCCCTGTCG  
 5761 GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC  
 5821 TATGGAAAAA CGCCAGCAAC CGGGCCTTTT TACGGTTCCCT GGCCTTTGC TGGCCTTTG  
 5881 CTCACATGTT TTTTCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG  
 5941 AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG  
 6001 AAGCGGAAGA GCGCCTGATG CGGTATTTC TCCCTACGCA TCTGTGCGGT ATTTCACACC  
 6061 GCAGACCAGC CGCGTAACT GGAAAATCG GTTACGGTT AGTAATAAAAT GGATGCCCTG  
 6121 CGTAAGCGGG TGTGGCGGA CAATAAGTC TTAAACTGAA CAAAATAGAT CTAAACTATG-

FIGURE 40C

110/240

6181 ACAATAAAAGT CTTAAACTAG ACAGAATAGT TGTAAAATG AATCAGTCCA GTTATGCTGT  
6241 GAAAAAGCAT ACTGGACTTT TGTTATGGCT AAAGCAAAC TTTCAATTTC TGAAGTGCAA  
6301 ATTGCCCGTC GTATTAAGA GGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCCGCGC  
6361 GTTGTGACAA TTTACCGAAC AACTCCGCGG CGGGGAAGCC GATCTCGGCT TGAACCGAATT  
6421 GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCATCAC TTCTTCCCCGT ATGCCCAACT  
6481 TTGTATAGAG ACCCACTGCG GGATCGTCAC CGTAACTCTGC TTGCACGTAG ATCACATAAG  
6541 CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG CGCCGTGGCA ATGCCCTGCC  
6601 TCCGGTGCTC GCCGGAGACT GCGAGATCAT AGATATAGAT CTCACTACCGC GGCTGCTCAA  
6661 ACCTGGGCAG AACGTAAGCC GCGAGAGCGC CAACAAACCGC TTCTTGGTCG AAGGCAGCAA  
6721 GCGCGATGAA TGTCCTTACTA CGGAGCAAGT TCCCCGAGGTA ATCGGAGTCC GGCTGATGTT  
6781 GGGAGTAGGT GGCTACGTCT CGAACTCAC GACCGAAAAG ATCAAGAGCA GCCCGCATGG  
6841 ATTTGACTTG GTCAAGGGCCG AGCCTACATG TGCGAAATGAT GCCCCACTT GAGCCACCTA  
6901 ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG  
6961 CTGCTCCATA ACATCAAACA TCGACCCACG GCGTAACGCG CTTCGCTGCTT GGATGCCCGA  
7021 GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA AAACCG

FIGURE 40D

111/240

Figure 4(A):

PDEST21

2-Hybrid Vector with  
DNA-Binding Domain

## ADH Promoter

700 ttg ccg ctc tgc tat caa gta taa ata gac ctg cca tta tta atc ttt tgt //  
 ,aac aac cda acg ata tat cat att tat ctg gac att aat aat tag aaa aca //

751 " ttc ctc gtc att gtt ctc gtt ccc ttt cct tgt ttc ttt ttc tgc aca //  
 ,aac aac cda caa caa gaa caa ggg aaa gaa gga aca aac aac aac acg tgt //

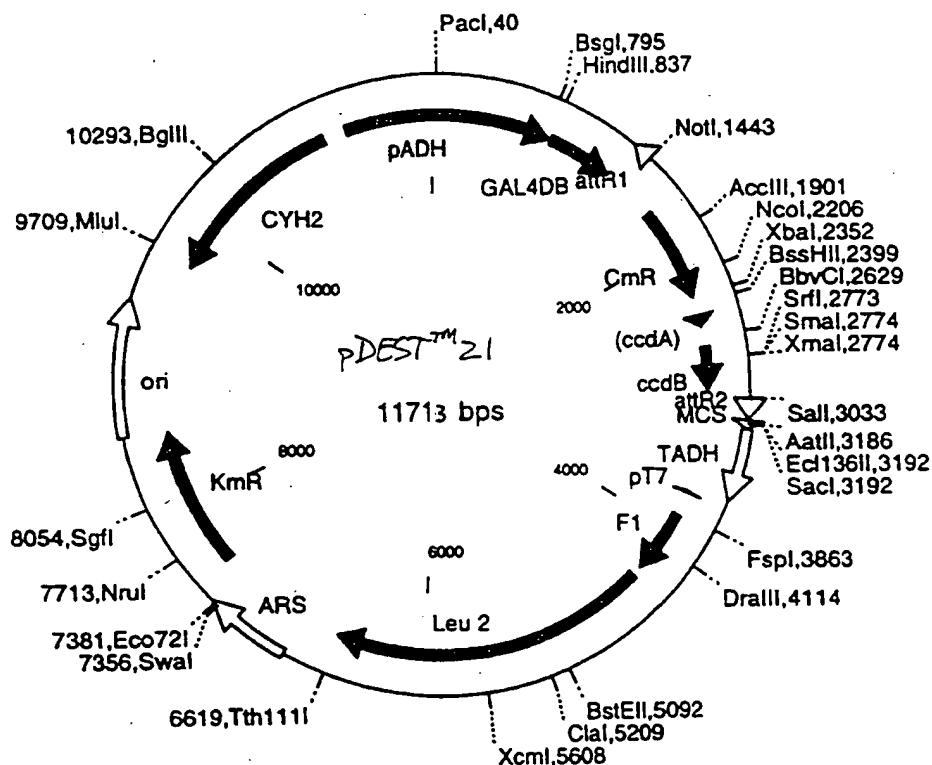
802 " ata ttt caa gct ata cca agc ata caa tca act cca agc ttg aag caa gcc //  
 ,tat aaa gtt cga tat gtt tcg tat gtt aat tga ggt tcg aac ttc gtt cgg  
 Start Transl M K L L S S Gal4-DE

853 tcc tga aag atg aag cta ctg tct tct atc gaa caa gca tgc gat att tgc //  
 agg act ttc tac ttc gat gac aga aga tag ctt gtt cgt acg cta taa acg //

1261 gaa gag agt agt aac aaa ggt caa aga cag ttg act gta tgc tgc agg tgc //  
 ctt ctc tca tca ttg ttt cca gtt tct gtc aac tga cat agc agc tcc agc  
 N Q T S L Y K K A att R1

1312 aat caa aca agt tng tac aaa aaa gct gaa cga gaa agc taa aat gat ata //  
 tca gtt tgt tca aac atg ttt ttt cga ctt gct ctt tgc att tta cta tat //

INT ↓



112/240

## pDEST21 11713 bp (rotated to position 11000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
857..1322	GAL4DB
1456..1332	attR1
1706..2365	CmR
2485..2569	inactivated ccdA
2707..3012	ccdB
3053..3177	attR2
3716..3735	pT7 (T7 promoter)
3899..4354	f1 (f1 intergenic region)
4414..6642	Leu2
7541..8515	kanR
9668..10958	CYH2
11118..848	pADH (ADH promoter)

1 TTTATTATGT TACAATATGG AAGGGAACCT TACACTTCTC CTATGCACAT ATATTAATTA  
 61 AAGTCCAATG CTAGTAGAGA AGGGGGGTAA CACCCCTCCG CGCTCTTTTC CGATTTTTT  
 121 CTAAACCGTG GAATATTCG GATATCCTT TGTTGTTCC GGGTGTACAA TATGGACTTC  
 181 CTCTTTCTG GCAACCAAAC CCATACATCG GGATTCTAT AATACCTTCG TTGGTCTCCC  
 241 TAACATGTAG GTGGCGGAGG GGAGATATAAC AATAGAACAG ATACCAGACAG AGACATAATG  
 301 GGCTAAACAA GACTACACCA ATTACACTGC CTCATTGATG GTGGTACATA ACGAACTAAT  
 361 ACTGTAGCCC TAGACTTGAT AGCCATCATC ATATCGAAGT TTCACTACCC TTTTTCCATT  
 421 TGCCATCTAT TGAAGTAATA ATAGGCGCAT GCAACTTCTT TTCTTTTTT TTCTTTCTC  
 481 TCTCCCCGT TGTGTCCTA CCATATCCGC AATGACAAAA AAAATGATGG AAGACACTAA  
 541 AGGAAAAAAAT TAACGACAAA GACAGCACCA ACAGATGTG TTGTTCCAGA GCTGATGAGG  
 601 GGTATCTTCG AACACACGAA ACTTTTCTC TCCTTCATTG ACGCACACTA CTCTCTAAATG  
 661 AGCAACGGTA TACGGCCTTC CTTCCAGTTA CTTGAATTG AAATAAAAAA AGTTTGGCGC  
 721 TTTGCTATCA AGTATAAATA GACCTGCAAT TATTAATCTT TTGTTTCTC GTCATTGTT  
 781 TCGTTCCCTT TCTTCCTGT TTCTTTCT GCACAATATT TCAAGCTATA CCAAGCATA  
 841 AATCAACTCC AAGCTTGAAG CAAGCCTCCT GAAAGATGAA GCTACTGTCT TCTATCGAAC  
 901 AAGCATGCGA TATTTGCCGA CTTAAAAAGC TCAAGTGCTC CAAAGAAAAA CCGAAGTGC  
 961 CCAAGTGTCT GAAGAACAAAC TGGGAGTGT GCTACTCTCC CAAAACCAAA AGGTCTCCGC  
 1021 TGACTAGGGC ACATCTGACA GAAGTGGAAAT CAAGGCTAGA AAGACTGGAA CAGCTATT  
 1081 TACTGATTTT TCCTCGAGAA GACCTTGACA TGATTTGAA AATGGATTCT TTACAGGATA  
 1141 TAAAAGCATT GTTAACAGGA TTATTTGTAC AAGATAATGT GAATAAAGAT GCCGTCA  
 1201 ATAGATTGGC TTCAGTGGAG ACTGATATGC CTCTAACATT GAGACAGCAT AGAATAAGTG  
 1261 CGACATCATC ATCGGAAGAG AGTAGTAACA AAGGCTAAAG ACAGTTGACT GTATCGTC  
 1321 GGTCGAATCA AACAAGTTTG TACAAAAAAG CTGAACGAGA AACGTAAT GATATAAATA  
 1381 TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAA  
 1441 ATATCCAGTC ACTATGGCGG CCGCTAAGTT GGCAGCATCA CCCGACGCAC TTTGCGCG  
 1501 ATAAATACCT GTGACGGAAG ATCACTTCG AGAATAAATA AATCCTGGTG TCCCTGTT  
 1561 TACCGGGAAAG CCCTGGGCCA ACTTTGGCG AAAATGAGAC GTTGATCGGC ACGTAAGAGG  
 1621 TTCCAACCTT CACCATATAATG AAATAAGATC ACTACCGGC GTATTTTTG AGTTATCG  
 1681 ATTTTCAGGA GCTAAGGAAG CTAAATGGA GAAAAAAATC ACTGGATATA CCACCGTT  
 1741 TATATCCCAA TGGCATCGTA AAGAACATT TGAGGCAATT CAGTCAGTT CTCATGTAC  
 1801 CTATAACCAAG ACCGTTCAAGC TGGATATTAC GGCCTTTTA AAGACCGTAA AGAAAAATAA  
 1861 GCACAAGTTT TATCCGGCT TTATTCACAT TCTTGGCCGC CTGATGAATG CTCATCCGG  
 1921 ATTCGGTATG GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGT  
 1981 CACCGTTTC CATGAGCAAA CTGAAACGTT TTCATCGCTC TGAGGTGAAT ACCACGACGA  
 2041 TTTCCGGCAG TTTCTACACA TATATTCGCA AGATGTGGCG TGTACGGTG AAAACCTGG  
 2101 CTATTTCCCT AAAGGGTTA TTGAGAATAT GTTTTCTGTC TCAGCCAATC CCTGGGTGAG  
 2161 TTTCACCAAGT TTTGATTTAA ACGGTGGCCAA TATGGACAAAC TTCTTCGCC CCGTTTC  
 2221 CATGGGCAAA TATTATACGC AAGGGGACAA GGTGCTGATG CCCTGGCGA TTCAGGTT  
 2281 TCATGCCGTC TGTGATGGCT TCCATGTCGG CAGAATGCTT AATGAATTAC AACAGTACTG  
 2341 CGATGAGTGG CAGGGCGGGG CGTAATCTAG AGGATCCGGC TTACTAAAAG CCAGATAACA  
 2401 GTATGCGTAT TTGCGCGCTG ATTTTGCGG TATAAGAATA TATACTGATA TGTATAACCG-

FIGURE 413

113/260

2461 AAGTATGTCA AAAAGAGGTG TGCTATGAAG CAGCGTATT AAGTGACAGT TGACAGCGAC  
 2521 AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA  
 2581 TGCAGAATGA AGCCCGTCGT CTGCGTCCG AACCGCTGGAA AGCGGAAAAT CAGGAAGGGA  
 2641 TGGCTGAGGT CGCCCGGGTT ATTGAAATGA ACGGCTCTT TGCTGACGAG AACAGGGACT  
 2701 GGTAAATGC AGTTAACGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTGTG  
 2761 GATGTACAGA GTGATATTAT TGACACCCCCC GGGCGACGGA TGTTGATCCC CCTGGCCAGT  
 2821 GCACGTCTGC TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCAGGGAT  
 2881 GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCAGGGAA  
 2941 GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATAAAAA ACGCCATTAA CCTGATGTT  
 3001 TGGGGAAATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAAGGTG ACCATAGTGA  
 3061 CTGGGATATGT TGTTGTTTAC AGTATTATGT AGTCTGTTT TTATGCAAAA TCTAATTAA  
 3121 TATATTGATA TTATATCAT TTACGTTTC TCAGTTCAGCT TTCTTGTACA AAGTGGTTG  
 3181 ATGGCCGCTA AGTAAGTAAG ACGTCGAGCT CTAAGTAAGT AACGGCCGCC ACCGCGGTGG  
 3241 AGCTTGGAC TTCTCGCCA GAGGTTGGT CAAGTCTCCA ATCAAGGTTG TCGGCTTGT  
 3301 TACCTGCCA GAAATTTACG AAAAGATGGA AAAGGGTCAA ATCGTTGGTA GATACGTTGT  
 3361 TGACACITCT AAATAAGCGA ATTTCTTATG ATTTATGATT TTATTTATTA AATAAGTTAT  
 3421 AAAAAAAATA AGTGTATACA AATTTAAAG TGACTCTTAG GTTTTAAAC GAAAATTCTT  
 3481 ATTCTTGAGT AACTCTTCC TGTAGGTCAAG TTGCTTTCT CAGGTATAGC ATGAGGTCGC  
 3541 TCTTATTGAC CACACCTCTA CCGGCATGCC GAGCAAATGC CTGCAAATCG CTCCCCATT  
 3601 CACCCAAATTG TAGATATGCT AACTCCAGCA ATGAGTTGAT GAATCTCGT GTGTATT  
 3661 TGTCCCTCAGA GGACAATACC TGTTGTAATC GTTCTTCCAC ACGGATCCA ATTGCGCC  
 3721 TAGTGAGTCG TATTACAATT CACTGGCGT CGTTTACAA CGTCGTGACT GGGAAAACCC  
 3781 TGGCGTTACC CAACTTAATC GCCTTGAGC ACATCCCCCT TTGCGCAGCT GGCAGTAATAG  
 3841 CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGAC  
 3901 GCGCCCTGTA GCGGCGCATT AAGCGCGCG GGTGTTGGTGG TTACGCGCAG CGTGCACCG  
 3961 ACACCTGCCA GCGCCCTAGC GCCCCTCCT TTGCTTTCT TCCCTCCCT TCTGCCACG  
 4021 TTGCGCCGCT TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC TTGTTAGGGT CCGATTTAGT  
 4081 GCTTACGGC ACCTCGACCC CAAAAAACTT GATTAGGGTG ATGGTTACAG TAGTGGGCC  
 4141 TCGCCCTGAT AGACGGTTT TCGCCCTTIG ACGTTGGAGT CCACGTTCTT TAATAGTGG  
 4201 CTCTTGTTC AACTGGAAC AACACTCAAC CCTATCTCGG TCTATTCTT TGATTATAA  
 4261 GGGATTTTGC CGATTTCGGC CTATTGTTA AAAATGAGC TGATTTAAC AAAATTTAAC  
 4321 GCGAATTTA ACAAAATATT AACGTTTACA ATTTCTGAT GCGGTATTCTT CTCCCTACGC  
 4381 ATCTGTGCGG TATTTCACAC CGCATATCGA CGCGTCGAGG AGAACCTCTA GTATATCCAC  
 4441 ATACCTAATA TTATTGCCTT ATTAAAATG GAATCGGAAC ATTACATCA AAATCCACAT  
 4501 TCTCTTCAAA ATCAATTGTC CTGTACTTCC TTGTTCATGT GTGTTCAAAA ACGTTATATT  
 4561 TATAGGATAA TTATACTCTA TTCTCAACA AGTAATTGGT TGTTGGCCG AGCGGTCTAA  
 4621 GGCCTCTGAT TCAAGAAATA TCTTGACCGC AGTTAACTGT GGGAAATACTC AGGTATCGTA  
 4681 AGATGCAAGA GTTCAATCT CTTAGCAACC ATTATTTTT TCCCTAACAT AACGAGAAC  
 4741 CACAGGGCG CTATCGCACA GAATCAAATT CGATGACTGG AAATTTTTG TTAATTTICAG  
 4801 AGGTCGCGCTG ACGCATATAC CTTTTCAAC TGAAAAATTG GGAGAAAAAG GAAAGGTGAG  
 4861 AGGCCGGAAC CGGTTTTCA TATAGAATAG AGAACGCGTC ATGACTAAAT GCTTGCATCA  
 4921 CAATACTTGA AGTTGACAAT ATTATTTAAG GACCTATTGT TTGTTCCAAT AGGTGGTTAG  
 4981 CAATCGTCTT ACTTTCTAAC TTTTCTTAC TTTTACATT CAGCAATATA TATATATAATT  
 5041 TCAAGGATAT ACCATTCTAA TGTCTGCC TATGCTGCTCC CCTAAGAAGA TCGTCGTTT  
 5101 GCCAGGTGAC CACGTTGGTC AAGAAATCAC AGCGGAAGCC ATTAAGGTT TAAAGCTAT  
 5161 TTCTGATGTT CGTTCCAATG TCAAGTTGCA TTGCAAATG CTTTAATTG GTGGTGTG  
 5221 TATCGATGCT ACAGGTGTCC CACTTCCAGA TGAGGGCGTG GAAGCTCCA AGAAGGTTGA  
 5281 TGCCGTTTG TTAGGTGCTG TGGGTGGTCC TAAATGGGGT ACCGGTAGTG TTAGACCTGA  
 5341 ACAAGGTTA CTAAAAATCC GTAAAGAACT TCAATTGTCAC GCCAACCTAA GACCATGTA  
 5401 CTTTGCATCC GACTCTTCTT TAGACTTATC TCCAAATCAAG CCACAAATTG CTAAGGTAC  
 5461 TGACTTCGTT TTGTCAGAG AATTAGTGGG AGGTATTTCAC TTGTTGAAAGA GAAAGGAAGA  
 5521 CGATGGTGAT GGTGTGCTT GGGATAGTGA ACAATACACC GTTCCAGAAG TGCAAAGAAT  
 5581 CACAAAGATG GCCGCTTCA TGGCCCTACCA ACATGAGCCA CCATTGCCA TTTGGTCC  
 5641 GGATAAAGCT AATGTTTGG CCTCTTCAAG ATTATGGAGA AAAACTGTGG AGGAAACCAT  
 5701 CAAGAACGAA TTCCCTACAT TGAAGGTTCA ACATCAATTG ATTGATTCTG CGGCACATGAT  
 5761 CCTAGTTAAG AACCCAACCC ACCTAAATGG TATTATAATC ACCAGCAACA TGTTTGGTGA  
 5821 TATCATCTCC GATGAAGCCT CCGTTATCCC AGGTTCCCTG GTGGTGTG CATCTGCGTC  
 5881 CTTGGCCTCT TTGCCAGACA AGAACACCGC ATTTGGTTTG TACGAACCAT GCCACGGTTC

FIGURE 41C

114/240

5941 TGCTCCAGAT TTGCCAAAGA ATAAGGTGA CCCTATCGCC ACTATCTTGT CTGCTGCAAT  
 6001 GATGTTGAAA TTGTCATTGA ACTTGCTGAGA AGAAGGTAAG GCCATTGAAG ATGCAGTTAA  
 6061 AAAGGTTTTG GATGCAGGTA TCAGAACTGG TGATTTAGGT GGTTCCAACA GTACCACCGA  
 6121 AGTCGGTGTGAT GCTGTCGCCG AAGAAGTTAA GAAAATCCTT GCTTAAAAAG ATTCTCTTTT  
 6181 TTTATGATAT TTGTACATAA ACTTTATAAA TGAAATTCAAT AATAGAAACG ACACGAAATT  
 6241 ACAAAATGGA ATATGTTCAT AGGGTAGACG AACTATATA CGCAATCTAC ATACATTTAT  
 6301 CAAGAAGGAG AAAAAGGAGG ATAGTAAAGG AATACAGGTA AGCAAATTGA TACTAATGGC  
 6361 TCAACGTGAT AAGGAAAAAG AATTGCACTT TAACATTAAT ATTGACAAGG AGGAGGGCAC  
 6421 CACACAAAAA GTTAGGTGTGTA ACAGAAAATC ATGAAACTAC GATTCCTAAT TTGATATTGG  
 6481 AGGATTTCTC CTAAAAAAA AAAAATACAA CAAATAAAA AACTCAATG ACCTGACCAT  
 6541 TTGATGGAGT TTAAGTCAT ACCTTCTGTA ACCATTCCC ATAATGGTGA AAGTCCCTC  
 6601 AAGAATTTTA CTCTGTCAGA AACGGCTTA CGACGTAGTC GATATGGTGC ACTCTCAGTA  
 6661 CAATCTGCTC TGATGCCGCA TAGTTAACG AGCCCCGACA CCCGCAACA CCCGCTGACG  
 6721 CGCCCTGACG GGCTTGTCTG CTCCCCGGCAT CCGCTTACAG ACAAGCTGTG ACCGTCTCCG  
 6781 GGAGCTGCAT GTGTCAGAGG TTTTCAACCGT CATCACCGAA ACGCGCAGA CGAAAGGGCC  
 6841 TCGTGATACG CCTATTTTA TAGGTTAATG TCATGATAAT AATGGTTCT TAGGACGGAT  
 6901 CGCTTGCTG TAACTTACAC GCGCTCGTA TCTTTTAATG ATGGAATAAT TTGGGAATT  
 6961 ACTCTGTGTT TATTTTATTT TATGTTTTGT ATTTGGATTT TAGAAAGTAA ATAAAGAAGG  
 7021 TAGAAGAGTT ACGGAATGAA GAAAAAAA TAAACAAAGG TTTAAAAAAT TTCAACAAAAA  
 7081 AGCGTACTTT ACATATATAT TTATTAGACA AGAAAAGCAG ATTAAATAGA TATACATTG  
 7141 ATTAACGATA AGTAAAATGT AAAATCACAG GATTTCTGTG TGTGGTCTC TACACAGACA  
 7201 AGATGAAACA ATTCCGGCATT AATACCTGAG AGCAGGAAGA GCAAGATAAA AGGTAGTATT  
 7261 TGTGCGCAT CCCCTAGAG TCTTTACAT CTTCGAAAAA CAAAAACTAT TTTTTCTTA  
 7321 ATTTCTTTTA TTACTTTCTA TTTTAAATT ATATATTAT ATTAAAAAAT TTAAATTATA  
 7381 ATTATTTTA TAGCACGTGA TGAAAAGGAC CCAGGTGGCA CTTTCGGGG AAATGTGCGC  
 7441 GGAACCCCTA TTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA  
 7501 TAACCTGTAT AAATGCTTCA ATAATCTGCA GCTCTGGCCG GTGCTCTAAA ATCTCTGATG  
 7561 TTACATTGCA CAAGATAAAA ATATATCATC ATGAAACAATA AAACGTCTG CTTACATAAA  
 7621 CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA ACGTCTTGCT GGAGGCCGCG  
 7681 ATTAATTCC AACATGGATG CTGATTATA TGGGTATAAA TGGGCTCGCG ATAATGTCGG  
 7741 GCAATCAGGT GCGACAATCT TTCGATTGTA TGGGAGGCCG GATGCGCCAG AGTTGTTCT  
 7801 GAAACATGGC AAAGGTAGCG TTGCCAATGTA TGTTACAGAT GAGATGGTCA GACTAAACTG  
 7861 GCTGACGGAA TTATGCTC TTCCGACCAT CAAGCATTTT ATCCGTAATC CTGATGATGC  
 7921 ATGGTTACTC ACCACTGCGA TCCGCGGGAA AACAGCATTC CAGGTATTAG AAGAATATCC  
 7981 TGATTCAAGGT GAAAATATTG TTGATGCGCT GGCAGTGTGCT CGCGCCGGT TGCATTGAT  
 8041 TCCTGTTGT AATTGTCCTT TTAACAGCGA TCGCGTATTT CGTCTCGCTC AGGCGCAATC  
 8101 ACGAATGAAT AACGGTTTGG TTGATGCGAG TGATTTGAT GACGAGCGTA ATGGCTGGCC  
 8161 TGTTGAACAA GTCTGGAAAG AAATGCATAC GCTTTGCCA TTCTCACCGG ATTCACTCGT  
 8221 CACTCATGGT GATTCTCAC TTGATAACCT TATTGAC GAGGGAAAT TAATAGGTG  
 8281 TATTGATGTT GGACGAGTCG GAATCGCAGA CCGATACCG GATCTTGCCA TCCTATGGAA  
 8341 CTGCCTCGGT GAGTTTCTC CTTCAATTACA GAAACGGCTT TTTCAAAAAT ATGGTATTGA  
 8401 TAATCTGTAT ATGAATAAAAT TGCACTTCA TTGATGCTC GATGAGTTT TCTAATCAGA  
 8461 ATTGGTTAAT TGTTGTAAC ACTGGCAGAG CATTACGCTG ACTTGACGGG ACGGCCATG  
 8521 ACCAAAATCC CTTAACGTGA GTTTCTGTC CACTGAGCGT CAGACCCCGT AGAAAAGATC  
 8581 AAAGGATCTT CTTGAGATCC TTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA  
 8641 CCACCGCTAC CAGCGGTGGT TTGTTGCGG GATCAAGAGC TACCAACTCT TTTTCCGAAG  
 8701 GTAACCTGGCT TCAGCAGAGC GCAGATACCA AATACGTCC TTCTAGTGTG GCCGTAGTTA  
 8761 GGCCACCACT TCAAGAACTC TGAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA  
 8821 CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGCTCTACCG GGTTGGACTC AAGACGATAG  
 8881 TTACCGGATA AGGGCGAGCG GTGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG  
 8941 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGT ACCATTGAGA AAGGCCACG  
 9001 CTTCCGAAG GGAGAAAGGC GGACAGGTAT CGCGTAAGCG GCAGGGTCCG AACAGGAGAG  
 9061 CGCACGAGGG AGCTTCCAGG GGGGAACGCC TGGTATCTTT ATAGTCTCTGT CGGGTTTCGC  
 9121 CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCCGAG CCTATGGAAA  
 9181 AACGCCAGCA ACGCGGCCCTT TTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG  
 9241 TTCTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT  
 9301 GATACCGCTC GCGCGAGCCG AACGACGGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA  
 9361 GAGGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCAATTA ATGCAGCTGG-

FIGURE 41D

115/260

9421 CACGACAGGT TTCCCCACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAC  
9481 CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCCTAT GTTGTGTGGA  
9541 ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTC  
9601 GGAATTAAACC CTCACTAAAG GGAACAAAAG CTGGTACCGA TCCCAGCTT TGCAAATTAA  
9661 AGCCTTCGAG CGTCCCAAAA CCTTCTCAAG CAAGGTTTC AGTATAATGT TACATCGTA  
9721 CACCGTCTG TACAGAAAAA AAAGAAAAAT TTGAAATATA AATAACGTT TTAATACTAA  
9781 CATAACTATA AAAAATAAA TAGGGACCTA GACTTCAGGT TGTCTAACCTC CTTCCCTTTTC  
9841 GGTTAGAGCG GATGTGGGG GAGGGCGTGA ATGTAAGCGT GACATAACTA ATTACATGAT  
9901 ATCGACAAAG GAAAAGGGGC CTGTTTACTC ACAGGTTTT TTCAAGTAGG TAATTAAGTC  
9961 GTTTCTGTCT TTTTCCCTCTC TCAACCCACC AAAGGCCATC TTGGTACTTT TTTTTTTTTT  
10021 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT  
10081 TTTTTTTTTT TTTTTTTTTT TCATAGAAAT AATACAGAAAG TAGATGTTGA ATTAGATTAA  
10141 ACTGAAGATA TATAATTAT TGGAAAATAC ATAGAGCTTT TTGTTGATGC GCTTAAGCGA  
10201 TCAATTCAAC AACACCACCA GCAGCTCTGA TTTTTCTTC AGCCAACCTG GAGACGAATC  
10261 TAGCTTGAC GATAACTGGA ACATTTGGAA TTCTACCCCT ACCCAAGATC TTACCGTAAC  
10321 CGGCTGCCAA AGTGTCAATA ACTGGAGCAG TTTCCTTAGA AGCAGATTTC AAGTATTGGT  
10381 CTCTCTTGTC TTCTGGGATC AATGTCCACA ATTGTCCAA GTTCAAGACT GGCTTCCAGA  
10441 AATGAGCTTG TTGTTGTGG AAGTATCTCA TACCAACCTT ACCGAAATAA CCTGGATGGT  
10501 ATTTATCCAT GTTAATTCTG TGTTGATGTT GACCACCGGC CATACTCTA CCACCGGGGT  
10561 GCTTTCTGTG CTTACCGATA CGACCTTAC CGGCTGAGAC GTGACCTCTG TGCTTTCTAG  
10621 TCTTAGTGAA TCTGGAAGGC ATTCTTGATT AGTTGGATGA TTGTTCTGGG ATTTAATGCA  
10681 AAAATCACTT AAGAAGGAAA ATCAACGGAG AAAGCAAACG CCATCTTAAA TATACGGGAT  
10741 ACAGATGAAA GGGTTTGAAC CTATCTGGAA AATAGCATTAA ACAAGCGAA AACTGCGAG  
10801 GAAAATTGTT TCGCTCTCTG CGGGCTATTAC ACGCGCCAGA GGAAAATAGG AAAAATAACA  
10861 GGGCATTAGA AAAATAATT TGATTTGGT AATGTGTGGG TCCCTGGTGA CAGATGTTAC  
10921 ATTGGTTACA GTACTCTTGT TTTTGCTGTG TTTTCGATG AATCTCCAAA ATGGTTGTTA  
10981 GCACATGGAA GAGTCACCGA TGCTAAGTTA TCTCTATGTA AGCTACGTGG CGTGACTTTT  
11041 GATGAAGCCG CACAAGAGAT ACAGGATTGG CAACTGCAAA TAGAATCTGG GGATCCCCC  
11101 TCGAGATCCG GGATCGAAGA AATGATGGTA AATGAAATAG GAAATCAAGG AGCATGAAGG  
11161 CAAAAGACAA ATATAAGGGT CGAACGAAAA ATAAAGTGAA AAGTGTGAT ATGATGTATT  
11221 TGGCTTGTGCG GCGCCGAAAA AACGAGTTA CGCAATTGCA CAATCATGCT GACTCTGTGG  
11281 CGGACCCGCG CTCTTGCCGG CCCGGCGATA ACGCTGGCG TGAGGCTGTG CCCGGCGGAG  
11341 TTTTTTGTGCG CTGCATTTTC CAAGGTTTAC CCTGCGCTAA GGGGCGAGAT TGGAGAAGCA  
11401 ATAAGAATGCG CGGTTGGGGT TGCGATGATG ACGACACCGA CAACTGGTGT CATTATTTAA  
11461 GTTGGCGAAA GAACCTGAGT GCATTTGCAA CATGAGTATA CTAGAAGAAT GAGCCAAGAC  
11521 TTGCGAGACG CGAGTTGCC GGTGGTGCAGA ACAATAGAGC GACCATGACC TTGAAGGTGA  
11581 GACGCGCATA ACCGCTAGAG TACTTTGAAG AGGAAACAGC AATAGGGTTG CTACCACTAT  
11641 AAATAGACAG GTACATACAA CACTGGAAAT GTTGTCTGT TTGAGTACGC TTTCAATTCA  
11701 TTTGGGTGTG CAC

FIGURE 416

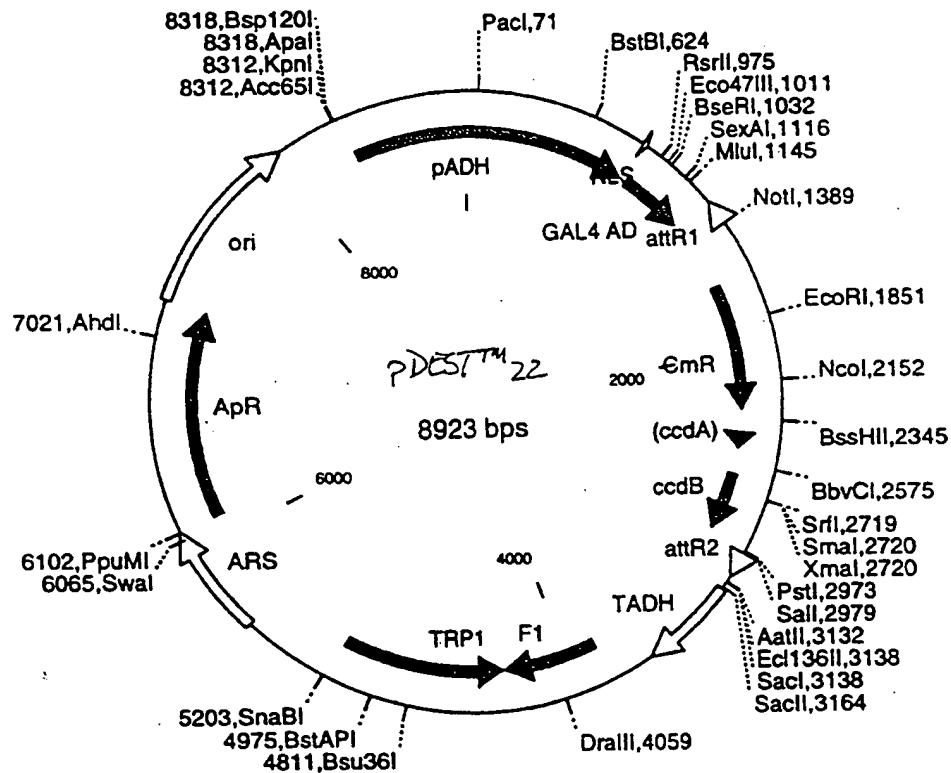
116/260

Figure 42A:

pDCST22

## 2-Hybrid Vector with Activation Domain

657 acg cac act act ctc taa tga gca acg gta tac ggc ctt cct tcc agt tac  
 tgc gtc tga tga gag att act cgt tgc cat atg ccg gaa gga agg tca atg  
 708 ttg aat ttg aaa taa aaa aag ttt gcc gct ttg cta tca agt ata aat aga  
 aac tta aac ttt att ttt ttc aaa cgg cga aac gat agt tca tat tta tct  
 759 cct gca att att aat ctt ttg ttt cct cgt cat tgt tct cgt tcc ctt tct  
 gga cgt taa taa tta gaa aac aaa gga gca gta aca aga gca agg gaa aga  
 810 //tcc/ttg/ttt/cct/ttg/cat/att/tca/agc/tat/acc/aag/cat/acg/atc//  
 ADH Promoter  
 861 //abc/tcc/aag/ctt/atg/ccc/aag/aag/cgg/aag/gtc/tcg/agc/ggc/gcc/aat//  
 Gal4-AD  
 //tgt/agg/ttc/gaa/tac/ggg/ttc/ttc/gcc/ttc/cag/agc/tcg/ccg/cgg/tca//  
 Start Translation  
 1218 gaa gat acc cca cca aac cca aaa aaa gag ggt ggg ttg aat cca [aca agt]  
 ctt cta tgg ggt ggt ttg ggt ttt ttt ctc cca ccc agc tta gtt [tgt tca]  
 1269 //L Y K K A attR1  
 //ttg/tac/aaa/aaa/gct/gaa/cga/gaa/acg/taa/a/  
 //aac/atg/ttt/ttt/cga/ctt/gct/ctt/tgc/att/t/  
 Intv



117/240

## pDEST22 8923 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
904..1248	GAL4 AD
1388..1264	attR1
1638..2297	CmR
2417..2501	inactivated ccdA
2639..2944	ccdB
2985..3109	attR2
3831..4318	f1 (f1 intergenic region)
4334..5176	TRP1
6110..7194	ampR
8344..866	pADH (yeast ADH promoter)

1 TTCATTTGGG TGTGCACTTT ATTATGTTAC AATATGGAAG GGAACCTTAC ACTTCCTCCTA  
 61 TGCACATATA TTAATTAAAG TCCAATGCTA GTAGAGAAGG GGGGTAACAC CCCTCCCGCG  
 121 TCTTTCCGA TTTTTTCTA AACCGTGGAA TATTTCGGAT ATCCTTTGT TGTTTCCGGG  
 181 TGTACAATAT GGACTTCCTC TTTTCTGGCA ACCAAACCCA TACATCGGGA TTCCCTATAAT  
 241 ACCTTCGTTG GTCTCCCTAA CATGTAGGTG GCGGAGGGGA GATATACAAT AGAACAGATA  
 301 CCAGACAAGA CATAATGGGC TAAACAAGAC TACACCAATT ACACTGCCTC ATTGATGGTG  
 361 GTACATAACG AACTAATACT GTAGCCCTAG ACTTGATAGC CATCATCATA TCGAAGTTTC  
 421 ACTACCCCTT TTCCATTTCG CATCTATTGA AGTAATAATA GGCGCATGCA ACTTCCTTTTC  
 481 TTTTTTTTC TTTTCTCTC CCCCCGTTGT TGTCTCACCA TATCCGAAT GACAAAAAAA  
 541 ATGATGGAAG ACACCAAAGG AAAAAATTAA CGACAAAGAC AGCACCAACA GATGTCGTTG  
 601 TTCCAGAGCT GATGAGGGGT ATCTTCGAAC ACACGAAACT TTTTCCCTCC TTCATTCACTG  
 661 CACACTACTC TCTAATGAGC AACGGTATAC GGCCTTCCCTT CCAGTTACTT GAATTTGAAA  
 721 TAAAAAAAGT TTGCGCTTT GCTATCAAGT ATAAATAGAC CTGCAATTAT TAATCTTTG  
 781 TTTCCTCGTC ATTTGTTCTCG TTCCCTTTCT TCCTGTTTC TTTTCTGCA CAATATTCA  
 841 AGCTATACCA AGCATACAAT CAACTCCAAG CTTATGCCA AGAAGAAGCG GAAGGTCTCG  
 901 AGCGCGCCA ATTTAATCA AAGTGGGAAT ATTGCTGATA GCTCATTTGTC CTTCACTTT  
 961 ACTAACAGTA GCAACGGTCC GAAACCTCATA ACAACTCAA CAAATTCTCA AGCGCTTTCA  
 1021 CAACCAATTG CCTCCTCTAA CGTTCATGAT AACTTCATGA ATAATGAAAT CACGGCTAGT  
 1081 AAAATTGATG ATGGTAATAA TTCAAAACCA CTGTCACCTG GTGGACGGG CCAAACATGCG  
 1141 TATAACCGT TTGGAATCAC TACAGGGATG TTTAATACCA CTACAATGGA TGATGTATAT  
 1201 AACTATCTAT TCGATGATGA AGATACCCCA CCAAACCCAA AAAAGAGGG TGGGTCGAAT  
 1261 CAAACAAGTT TGTACAAAAA AGCTGAACGA GAAACGTAAA ATGATATAAA TATCAATATA  
 1321 TTAAATTAGA TTTTGCATAA AAAACAGACT ACATAATACT GTAAAACACA ACATATCCAG  
 1381 TCACTATGGC GGGCGCTAAG TTGGCAGCAT CACCCGACGC ACTTTGCGCC GAATAAAATAC  
 1441 CTGTGACGGA AGATCACTTC GCAGAATAAA TAAATCCTGG TGTCCCTGTT GATACGGGA  
 1501 AGCCCTGGGC CAACTTTTGG CGAAAATGAG ACGTTGATCG GCACGTAAGA GGTTCCAAGT  
 1561 TTCACCATAA TGAAATAAGA TCACTACCGG GCGTATTTTG TGAGTTATCG AGATTTTCAG  
 1621 GAGCTAAGGA AGCTAAAATG GAGAAAAAAA TCACTGGATA TACCCACCGTT GATATATCCC  
 1681 AATGGCATCG TAAAGAACAT TTGAGGCAT TTCAGTCAGT TGCTCAATGT ACCTATAACC  
 1741 AGACCGTTCA GCTGGATATT ACGGCCTTT TAAAGACCGT AAAGAAAAAT AAGCACAAGT  
 1801 TTTATCCGGC CTTTATTAC C ATTCTTGCCC GCCTGATGAA TGCTCATCCG GAATTCCGTA  
 1861 TGGCAATGAA AGACGGTGAG CTGGTGTAT GGGATAGTGT TCACCCCTTGT TACACCGTTT  
 1921 TCCATGAGCA AACTGAAACG TTTTCACTCG TCTGGAGTGA ATACCAACGAC GATTTCCGGC  
 1981 AGTTTCTACA CATATATTGCA CAAGATGTTGG CGTGTACGG TGAAAACCTG GCCTATTTC  
 2041 CTAAAGGGTT TATTGAGAAT ATGTTTTTCG TCTCAGCCAA TCCCTGGGTG AGTTTCAACCA  
 2101 GTTTTGATTAA AACATGGGCA AATATGGACA ACTTCTTCGC CCCCGTTTTC ACCATGGGCA  
 2161 AATATTATAC GCAAGGGCAG AAGGTGCTGA TGCCGCTGGC GATTCAAGGTT CATCATGCC  
 2221 TCTGTGATGG CTTCCATGTC GGCAGAAATGC TTAATGAATT ACAACAGTAC TGCGATGAGT  
 2281 GGCAGGGCGG GGCAGTAATCT AGAGGATCCG GCTTACTAAA AGCCAGATAA CAGTATGCGT  
 2341 ATTTGCGCGC TGATTTTTGCA GGTATAAGAA TATATACTGA TATGTATACC CGAAGTATGT  
 2401 CAAAAAGAGG TGTGCTATGA AGCAGCGTAT TACAGTGACA GTTGACAGCG ACAGCTATCA  
 2461 GTTGCTCAAG GCATATATAC TGTCAATATC TCCGGCTGG TAAGCACAAC CATGCAGAAT  
 2521 GAAGCCCGTC GTCTGCGTGC CGAACGCTGG AAAGCGGAAA ATCAGGAAGG GATGGCTGAG-

FIGURE 425

118/240

2581 GTCGCCCGGT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACAGGGG CTGGTGAAAT  
 2641 GCAGTTTAAG GTTACACCT ATAAAAGAGA GAGCCGTTAT CGTCTGTTG TGGATGTACA  
 2701 GAGTGATATT ATTGACACGC CCGGGCGACG GATGGTGTAC CCCCTGGCCA GTGCACGTCT  
 2761 GCTGTCAGAT AAAGTCTCCC GTGAACCTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG  
 2821 GCGCATGATG ACCACCGATA TGGCCAGTGT GCCGGTCTCC GTTATCGGGG AAGAAGTGGC  
 2881 TGATCTCAGC CACCGCGAAA ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGAAT  
 2941 ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT GACTGGATAT  
 3001 GTTGTGTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATT AATATATTGA  
 3061 TATTATATC ATTACGTT TCTCGTCAG CTTCTTGTA CAAAGTGGTT TGATGCCGC  
 3121 TAAGTAAGTA AGACGTGAG CTCTAAGTAA GTAACGGCCG CCACCGCGGT GGAGCTTGG  
 3181 ACTTCTTCGC CAGAGGTTG GTCAAGTCTC CAATCAAGGT TGTCGGCTTG TCTACCTTGC  
 3241 CAGAAATTAA CGAAAAGATG GAAAAGGGTC AAATCGTTGG TAGATACGTT GTTGACACTT  
 3301 CTAAATAAGC GAATTTCTTA TGATTTATGA TTTTATTAT TAAATAAGTT ATAAAAAAA  
 3361 TAAGTGTATA CAAATTTAA AGTGAACCTT AGGTTTTAAA ACGAAAATTC TTATTCCTGA  
 3421 GTAACCTTT CCTGTAGGTC AGGTTGCTT CTCAGGTATA GCATGAGGTC GCTCTTATIG  
 3481 ACCACACCTC TACCGGCATG CCGAGCAAAT GCCTGCAAAT CGCTCCCCAT TTCACCCAAT  
 3541 TGTAGATATG CTAACCTCAG CAATGAGTTG ATGAATCTCG GTGTGTATT TATGTCCTCA  
 3601 GAGGACAATA CCTGTTGTAA TCGTTCTCC ACACGGATCC CAATTCGCC C TATAGTGAGT  
 3661 CGTATTACAA TTCACTGGCC GTGTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA  
 3721 CCCAACCTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG  
 3781 CCCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG ACGCCCTG  
 3841 TAGCGCGCA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG CTACACTTGC  
 3901 CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTGCCA CGTTGCCCGG  
 3961 CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT CCCTTITAGGG TTCCGATTTA GTGCTTACG  
 4021 GCACCTCGAC CCCAAAAAAAC TTGATTAGGG TGATGGTTCA CTGATGGGC CATGCCCTG  
 4081 ATAGACGGTT TTTCGCCCC TTGACGTTGGA GTCCACGTT TTTAATAGT GACTCTTGT  
 4141 CCAAACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTGATTAT AAGGGATTTT  
 4201 GCGGATTTCG GCCTATTGGT TAAAAAATGA GCTGATTAA CAAATTTA ACGCAGAATT  
 4261 TAACAAAATA TTAACGTTA CAATTTCTG ATGCGGTATT TTCTCTTAC GCATCTGTGC  
 4321 GGTATTTCAC ACCGCAGGCC AGTGCACAAA CAATACTTAA ATAATACTA CTCAGTAATA  
 4381 ACCTATTTCT TAGCATTTT GACGAAATT GCTATTGT TAGAGTCTTT TACACCATT  
 4441 GTCTCCACAC CTCCGCTTAC ATCAACACCA ATAACGCCAT TTAATCTAAG CGCATCACCA  
 4501 ACATTTCTG GCGTCAGTCC ACCAGCTAAC ATAAAATGTA AGCTTTCGGG GCTCTTGC  
 4561 CTTCCAACCC AGTCAGAAAT CGAGTTCAA TCCAAAAGTT CACCTGTCCC ACCTGTTCT  
 4621 GAATCAAACA AGGGAATAAA CGAATGAGGT TTCTGTGAAG CTGCACTGAG TAGTATGTG  
 4681 CAGTCTTTG GAAATACGAG TCTTTAATA ACTGGCAAAC CGAGGAACCTC TTGGTATTCT  
 4741 TGCCACGACT CATCTCCATG CAGTTGGACG ATATCAATGC CGTAATCATT GACCAGAGCC  
 4801 AAAACATCCT CTTTAGGTTG ATTACGAAAC ACGCCAACCA AGTATTTCGG AGTGCCTGAA  
 4861 CTATTTTAT ATGCTTTAC AAGACTTGAA ATTTCTTGC CAATAACCGG GTCAATTGTT  
 4921 CTCTTCTAT TGGCACACA TATAATACCC AGCAAGTCAG CATCGGAATC TAGAGCACAT  
 4981 TCTGCGCCT CTGTGCTCTG CAAGCCCAA ACTTTACCA ATGGACCAGA ACTACCTGTG  
 5041 AAATTAATAA CAGACATACT CCAAGCTGCC TTTGTGTCT TAATCACGTA TACTCACGTG  
 5101 CTCAAAGTC ACCAATGCC TCCCTCTTGG CCCTCTCCCT TTCTTTTTC GACCAGATTA  
 5161 ATTCTTAATC GGCAAAAAAA GAAAAGCTCC GGATCAAGAT TGTACGTAAG GTGACAAGCT  
 5221 ATTTTTCAAT AAAGAATATC TTCCACTACT GCCATCTGGC GTCATAACTG CAAAGTACAC  
 5281 ATATATTACG ATGCTGTCTA TTAAATGCTT CCTATATTAT ATATATAGTA ATGTCCTTAA  
 5341 TGGTCACTC TCAGTACAAT CTGCTCTGAT GCGCATAGT TAAGCCAGCC CCGACACCCG  
 5401 CCAACACCCG CTGACGCGCC CTGACGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA  
 5461 GCTGTGACCG TCTCCGGAG CTGCGATGTGT CAGAGGTTT CACCGTCATC ACCGAAACGC  
 5521 GCGAGACGAA AGGGCCTCGT GATACGCCA TTTTATAGG TTAATGTCAT GATAATAATG  
 5581 GTTTCTTAGG ACGGATCGCT TGCCTGTAAC TTACACGCGC CTCGTATCTT TTAATGATGG  
 5641 AATAATTGG GAATTTACTC TGTGTTTATT TATTTTATG TTTTGTTATT GGATTTAGA  
 5701 AAGTAAATAA AGAAGGTAGA AGAGTTACGG AATGAAGAAA AAAAAATAAA CAAAGTTTA  
 5761 AAAAATTTCAC ACACAAAGCG TACTTTACAT ATATATTTAT TAGACAAGAA AAGCAGATTA  
 5821 AATAGATATA CATTGATTAA ACGATAAGTA AAATGTAAAA TCACAGGATT TTCGTGTG  
 5881 GTCTTCTACA CAGACAAGAT GAAACAATTC GGCATTAATA CCTGAGAGCA GGAAGAGCAA  
 5941 GATAAAAGGT AGTATTGTT GGCGATCCCC CTAGAGTCTT TTACATCTC GGAAAACAAA  
 6001 AACTATTTT TCTTTAATT TTTCTATT TTTCTATT TAAATTTAT ATTTATATTA-

FIGURE 42c

119/240

6061 AAAAATTTAA ATTATAATTAA TTTTTATAGC ACGTGATGAA AAGGACCCAG GTGGCACTTT  
 6121 TCGGGGAAAT GTGCGCGGAA CCCCTATTG TTTATTTTC TAAATACATT CAAATATGTA  
 6181 TCCGCTCATG AGACAATAAC CCTGATAAAAT GCTTCATAA TATTGAAAAAA GGAAGAGTAT  
 6241 GAGTATTCAA CATTTCGTG TCGCCCTTAT TCCCCTTTT GCGGCATTG GCCTTCCTGT  
 6301 TTTTGCTCAC CCAGAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG  
 6361 AGTGGGTTAC ATCGAAGCTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCGA  
 6421 AGAACGTTT CCAATGATGA GCACCTTAA AGTTCTGCTA TGTGGCGGG TATTATCCCG  
 6481 TATTGACGCC GGGCAAGAGC AACTCGGTG CCGCATAACAC TATTCTCAGA ATGACTTGGT  
 6541 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG  
 6601 CAGTGCCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTCTGCA CAACGATCGG  
 6661 AGGACCGAAG GAGCTAACCG CTTTTTTCA CAACATGGGG GATCATGTAA CTCGCCTTGA  
 6721 TCGTTGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC  
 6781 TGTAGCAATG GCAACAACGT TGCGCAAACCT ATTAACCTGGC GAACTACTTA CTCTAGCTTC  
 6841 CCGGCAACAA TTATAGACT GGATGGAGGC GGATAAAAGTT GCAGGACAC TTCTGCGCTC  
 6901 GGCCCCCTCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGGC GTGGGTCTCG  
 6961 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC  
 7021 GACGGGCAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC  
 7081 ACTGATTAAG CATTGGTAAC TGTCAAGACCA AGTTTACTCA TATATACTTT AGATTGATT  
 7141 AAAACTTCAT TTTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC  
 7201 CAAAATCCCT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA  
 7261 AGGATCTTCT TGAGATCCCTT TTTTCTGCG CGTAATCTGC TGTTGCAA CAAAAAAACC  
 7321 ACCGCTACCA GCGGTGGTTT GTTGCCTGGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT  
 7381 AACTGGCTTC AGCAGAGCG AGATACAAA TACTGTCTT CTAGTGTAGC CGTAGTTAGG  
 7441 CCACCACTTC AAGAACTCTG TAGCACCCTC TACATACCTC GCTCTGCTAA TCCTGTTACC  
 7501 AGTGGCTGCT GCGAGTGGC ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT  
 7561 ACCGGATAAG GCGCAGCGG CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA  
 7621 GCGAACGACC TACACCGAAC TGAGATAACCT ACAGCGTGG CATTGAGAAA GCGCCACGCT  
 7681 TCCCGAAGGG AGAAAGGCAG ACAGGTATCC GGTAAAGCGGC AGGGTCGGAA CAGGAGAGCG  
 7741 CACGAGGGAG CTTCCAGGGG GGAACGCCG GTATCTTTAT AGTCTGTG TGTTTCGCCA  
 7801 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCCGAGCC TATGGAAAAA  
 7861 CGCCAGCAAC GCGGCCCTT TACGGTTCCCT GGCCTTTGC TGGCCTTTG CTCACATGTT  
 7921 CTTTCCTGCG TTATCCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAAGCTGA  
 7981 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTC GTGAGCGAGG AAGCGGAAGA  
 8041 GCGCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCAATTAA GCAGCTGGCA  
 8101 CGACAGGTTT CCGACTGGA AAGCGGGCAG TGAGCGAAC GCAATTAAAG TGAGTTACCT  
 8161 CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG GCTCCTATGT TGTGTGGAAT  
 8221 TGTGAGCGGA TAACAATTTC ACACAGGAA CAGCTATGAC CATGATTACG CCAAGCTCGG  
 8281 AATTAAACCT CACTAAAGGG AACAAAAGCT GGGTACCGGG CCCCCCTCG AGATCCGGGA  
 8341 TCGAAGAAAT GATGTAAT GAAATAGGAA ATCAAGGAGC ATGAAGGCAA AAGACAAATA  
 8401 TAAGGGTCGA AGAAAATA AAGTAAAAG TGTGATATG ATGTATTG CTTTGCAGCG  
 8461 CCGAAAAAAAC GAGTTTACCG AATTGACAA TCATGCTGAC TCTGTGGCG ACCCGCGCTC  
 8521 TTGCCGGCCC GGCATAACG CTGGCGTGA GGCCTGTGCC GGCAGGAGTTT TTTGCCCTG  
 8581 CATTTCCAA GGTTCACCT GCGCTAAGGG GCGAGATTGG AGAAGCAATA AGAATGCCGG  
 8641 TTGGGGTTGC GATGATGACG ACCACGACAA CTGGTGTGAT TATTAAAGTT GCCGAAAGAA  
 8701 CCTGAGTGCA TTTGCAACAT GAGTACTA GAAGAATGAG CCAAGACTTG CGAGACGCGA  
 8761 GTTGTGCCGGT GGTGCGAACAA ATAGAGCGAC CATGACCTTG AAGGTGAGAC GCGCATAACC  
 8821 GCTAGAGTAC TTTGAAGAGG AAACAGCAAT AGGGTTGCTA CCAGTATAAA TAGACAGGTA  
 8881 CATAAACAC TGGAAATGGT TGTCTGTTG AGTACGCTTT CAA

FIGURE 4<sup>2D</sup>

120/240

pDEST23

## His6 carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA

205 atc ccg cga aat taa tac gac tca cta tag gga gat cac aac ggt ttc cct  
tag ggc gct tta att atg ctg agt gat acc cgt ctg gtg ttg cca aag gga  
attR1

256 cta gat cac aag ttt gta caa aaa agc tga acg aga aac gta aaa tga tat  
gat cta ctg ttc aaa cat gtt ttt tcg act tgc tct ttg cat ttt act ata //

1 — Cm<sup>R</sup> — ccd B — 11

1888 ttt tta tgc aaa atc taa ttt aat ata ttg ata ttt ata tca ttt tao gtt  
aaa sat acg ttt tag att aaa tta tat aac tat aaa tat agt aaa atg caa  
attR2 A F L Y K V Y I M S Y Y H H  
1939 tct cgt tca gct ttd ttg tac aaa gtg gtg att atg teg tac tac cat cac  
aga gca agt cga aag aac atg ttt cac cac taa tac aac atg atg atg gta gtg  
H H H H L D E V Q term His6 //  
1990 cat cac cat cac ctc gat gag caa taa cta gca taa ccc ctt ggg gec tet  
gta gtg gta gtg gag cta ctc gtt att gat cgt att ggg gaa ccc cgg aga

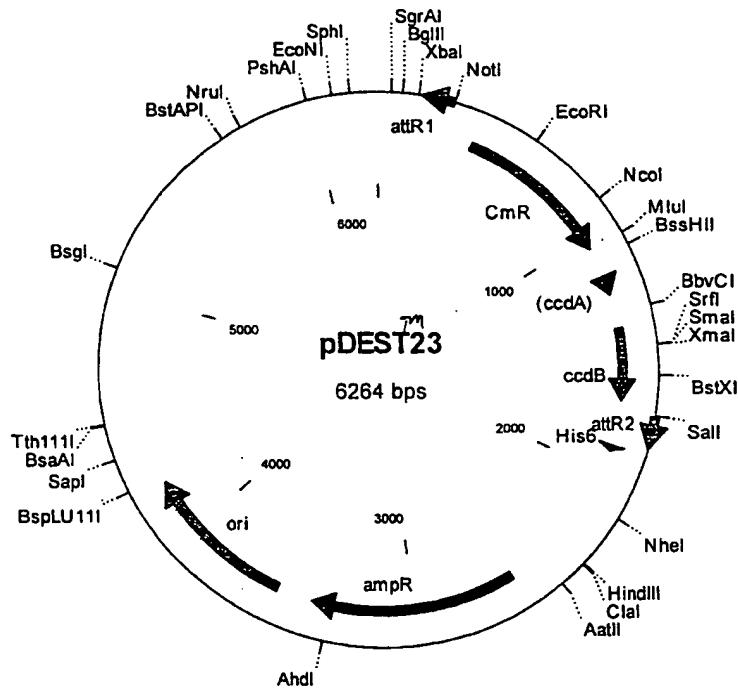


FIGURE 43A

121/240

## pDEST23 6264 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
285..161	attR1
394..1053	CmR
1173..1257	inactivated ccdA
1395..1700	ccdB
1741..1865	attR2
1883..1911	his6
2574..3434	ampR
3583..4222	ori

1 TCTTCCCCAT CGGTGATGTC GGCGATATAG GCGCCAGCAA CCGCACCTGT GGCGCCGGTG  
 61 ATGCCGGCCA CGATGCGTCC GGCGTAGAGG ATCGAGATCT CGATCCCGCG AAAATAATAC  
 121 GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGATC ACAAGTTGT ACAAAAAAGC  
 181 TGAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTAA ATTAGATTT TGCAATAAAA  
 241 ACAGACTACA TAATACTGTA AAACACAACA TATCAGTCA CTATGGCGGC CGCATTTAGGC  
 301 ACCCCCAGGCT TTACACTTTA TGCTTCCGGC TCGTATAATG TGTGGATTTT GAGTTAGGAT  
 361 CCGGCCAGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAAATCAC TGGATATACC  
 421 ACCGTTGATA TATCCAATG GCATCGTAA GAACATTTG AGGCATTTCA GTCAAGTTGCT  
 481 CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG CCTTTTTAAA GACCGTAAAG  
 541 AAAAATAAGC ACAAGTTTTA TCCGGCCTTT ATTACACATTC TTGCCCCGCT GATGAATGCT  
 601 CATCCCGAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA TAGTGTTCAC  
 661 CCTTGTACCA CCGTTTTCCA TGAGCAAAC GAAACGTTT CATCGCTCTG GAGTGAATAC  
 721 CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG ATGTGGCGTG TTACGGTGAA  
 781 AACCTGGCCT ATTCCCTAA AGGGTTTATT GAGAATATGT TTTTCGTCTC AGCCAAATCCC  
 841 TGGGTGAGTT TCACCAAGTT TGATTTAAC GTGGCCAATA TGGACAACCTT CTTCCGCCCC  
 901 GTTTTCACCA TGGGCAAATA TTATACGCAA GGCACAAAGG TGCTGATGCC GCTGGCGATT  
 961 CAGGTTCATC ATGCCGTCTG TGATGGCTTC CATGTCGGCA GAATGCTTAA TGAATTACAA  
 1021 CAGTACTGCG ATGAGTGGCA GGGCGGGCG TAAACGCGTG GATCCGGCTT ACTAAAAGCC  
 1081 AGATAACAGT ATGGCTATTG GCGCGCTGAT TTTTGCGGTAA TAAAGAATATA TACTGATATG  
 1141 TATACCGAA GTATGTCAAA AAGAGGTGTC CTATGAAGCA GCGTATTACA GTGACAGTTG  
 1201 ACAGCGACAG CTATCAGTTG CTCAAGGCAT ATATGATGTC AATATCTCCG GTCTGGTAAG  
 1261 CACAACCATG CAGAATGAAG CCCGTCGTCT GCGTGGCGAA CGCTGGAAAG CGGAAATCA  
 1321 GGAAGGGATG GCTGAGGTGCG CCCGGTTTAT TGAAATGAAC GGCTCTTTTG CTGACCGAGAA  
 1381 CAGGGACTGG TGAATGCAAG TTAAAGGTTT ACACCTATAA AAGAGAGAGC CGTTATCGTC  
 1441 TGTTTGTTGGA TGTCAGAGT GATATTATTG ACACGCCCGG GCGACGGATG GTGATCCCC  
 1501 TGGCCAGTGC ACGTCTGCTG TCAGATAAAAG TCTCCCGTGA ACTTTACCCG GTGGTGCATA  
 1561 TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC CAGTGTGCCCG GTCTCCGTAA  
 1621 TCGGGGAAGA AGTGGCTGAT CTCAGCCACC GCGAAAATGA CATAAAAAAAC GCCATTAACC  
 1681 TGATGTTCTG GGGAAATATAA ATGTCAGGCT CCCTTATACA CAGCCAGTCT GCAGGTCGAC  
 1741 CATAGTGACT GGATATGTTG TGTTTACAG TATTATGTTG TCTGTTTTT ATGCAAAATC  
 1801 TAATTTAATA TATTGATATT TATATCATTT TACGTTCTC GTTCAAGCTTTT CTTGTACAAA  
 1861 GTGGTGATTA TGTCGTACTA CCATCACCAT CACCATCACC TGATGAGCA ATAACATAGCA  
 1921 TAACCCCTTG GGGCCTCTAA ACGGGTCTTG AGGGGTTTT TGCTGAAAGG AGGAACATATA  
 1981 TCCGGATATC CACAGGACGG GTGTGGTCGC CATGATCGCG TAGTCGATAG TGGCTCCAAG  
 2041 TAGCGAAGCG AGCAGGACTG GGCAGCGGCC AAAGCGGTGCG GACAGTGCTC CGAGAACGGG  
 2101 TCGCATAGA AATTGATCA ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGGCGAT  
 2161 GCTGTCGGAA TGAGCATAT CCCGCAAGAG GCGCCGGCAGT ACCGGCATAA CCAAGCCTAT  
 2221 GCCTACAGCA TCCAGGGTGA CGGTGCCAG GATGACGATG AGCGCATTGT TAGATTTCAT  
 2281 ACACGGTGC TGACTGCGTT AGCAATTAA CTGTGATAAA CTACCGCATT AAAGCTTATC  
 2341 GATGATAAGC TGTCACACAT GAGAATTCTT GAAGACGAAA GGGCCTCGTG ATACGCCATT  
 2401 TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG  
 2461 GAAATGTGCG CGGAACCCCT ATTGTTTAT TTTCTAAAT ACATTCAAAT ATGTATCCGC  
 2521 TCATGAGACA ATAACCCCTGA TAAATGCTTC AATAATATTG AAAAGGAAG AGTATGAGTA  
 2581 TTCAACATTT CGGTGTCGCC CTTATTCCCT TTTTGCAGGC ATTTTGCCTT CCTGTTTTG  
 2641 CTCACCCAGA AACGCTGGTG AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG-

FIGURE 43B

122/260

2701 GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTCGC CCCGAAGAAC  
 2761 GTTTTCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCCTGTTG  
 2821 ACGCCGGCA AGAGCAACTC GGTGCCGCA TACACTATT TCAGAATGAC TTGGTTGAGT  
 2881 ACTCACCACTG CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGAGTG  
 2941 CTGCCATAAC CATGAGTGT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC  
 3001 CGAAGGAGCT AACCGCTTTT TTGACAAACA TGGGGATCA TGTAACCTCC CTTGATCGTT  
 3061 GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCAACG ATGCCCTGCAG  
 3121 CAATGGCAAC AACGTTGCGC AAACATTAA CTGGCGAACT ACCTACTCTA GCTTCCCGGC  
 3181 AACAAATTAAAT AGACTGGATG GAGGCGGATA AAGTGTGAGG ACCACTTCTG CGCTCGGCC  
 3241 TTCCGGCTGG CTGGTTTATT GCTGATAAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGTA  
 3301 TCATTGCAAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG  
 3361 GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA  
 3421 TTAAGCATTG GTAACGTGCA GACCAAGTTT ACTCATATAT ACCTTAGATT GATTTAAAAC  
 3481 TTCATTTTA ATTAAAAAGG ATCTAGGTGA AGATCCTTT TGATAATCTC ATGACCAAAA  
 3541 TCCCTTAACG TGAGTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT  
 3601 CTTCTTGAGA TCCTTTTTT CTGCGCTAA TCTGCTGCTT GCAAACAAAA AAACCAACGC  
 3661 TACCAGCGGT GGGTTGTTTG CCGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAACGT  
 3721 GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTCTAGT STAGCCGTAG TTAGGCCACC  
 3781 ACTTCAGAA CTCTGTAGCA CGCCTACAT ACCTCGCTC GCTAATCTC TTACCACTGG  
 3841 CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG  
 3901 ATAAGGCGCA GCGGTGGGG TGAAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA  
 3961 CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG  
 4021 AAGGGAGAAA GCGGGACAGG TATCCGGTAA CGGGCAGGGT CGGAACAGGA GAGCGCACGA  
 4081 GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTATAGTCC TGTCGGGTTT CGCCACCTCT  
 4141 GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG AAAAACGCCA  
 4201 GCAACCGCGC CTTTTACGG TTCCGGCCTT TTGCTGGCC TTTGCTCAC ATGTTCTTTC  
 4261 CTGCGTTATC CCTGATTCT GTGGATAACC GTATTACCGC TTGAGTGAGTGA GCTGATAACCG  
 4321 CTCGCCGCAG CGAACGACCG GAGCGCAGCG AGTCAGTGA CGAGGAAGCG GAAGAGCGCC  
 4381 TGATGCGGTA TTTCTCCTT ACGCATCTGT CGGGTATTTT ACACCGCATA TATGGTGCAC  
 4441 TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATACTCTC GCTATCGCTA  
 4501 CGTGACTGGG TCATGGCTGC GCCCCACAC CGCCTAACAC CGCTGACGG GCCCTGACGG  
 4561 GCTTGCTGC TCCCAGCATE CGCTTACAGA CAAGCTGTGA CGCTCTCCGG GAGCTGCATG  
 4621 TGTCAGAGGT TTTCACCGTC ATCACGAAA CGCGCGAGGC AGCTGCGGTA AAGCTCATCA  
 4681 GCGTGGTCGT GAAGCGATTC ACAGATGTCT GCCTGTTCAT CGCGTCCAG CTCGTTGAGT  
 4741 TTCTCCAGAA GCGTTAATGT CTGGCTCTG ATAAAGCGGG CCATGTTAAG GGGGGTTTTT  
 4801 TCCTGTTGG TCACTGATGC CTCCGTGAA GGGGATTTC TGTTCATGGG GGTAATGATA  
 4861 CCGATGAAAC GAGAGAGGAT GCTCACGATA CGGGTTACTG ATGATGAACA TGCCCCGGTTA  
 4921 CTGGAACGTT GTGAGGGTAA ACAACTGGCG GTATGGATGC GGCAGGGACCA GAGAAAAATC  
 4981 ACTCAGGGTC AATGCCAGCG CTTCGTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG  
 5041 CAGCATCCCTG CGATGCAGAT CGGAACATA ATGGTGCAGG GCGCTGACTT CGCGTCTTCC  
 5101 AGACTTTACG AAACACGGAA ACCGAAGACC ATTCACTGTT TGCTCAGGT CGCAGACGTT  
 5161 TTGCAGCAGC AGTCGCTTCA CGTTCGCTCG CGTATCGGTG ATTCAATTCTG CTAACCCAGTA  
 5221 AGGCAACCCC GCCAGCCTAG CGGGTCTC AACGACAGGA GCACGATCAT GCGCACCCGT  
 5281 GGCCAGGACC CAACGCTGCC CGAGATGCGC CGCGTGCAGG TGCTGGAGAT GGCAGACGCC  
 5341 ATGGATATGT TCTGCCAAGG GTGGTTTGC GAAGTTAGGC TGTTAAGAGC CGCGAGCGAT CCTTGAAGCT  
 5401 GCTCCAATTG TTGGAGTGGT GAATCCGTTA GCGAGGTGCC GCGGGCTTCC ATTCAAGTCG  
 5461 AGGTGGCCCG GCTCCATGCA CGCGACCGA ACAGGGGGAG GCAGACAAGG TATAGGGCGG  
 5521 CGCTTACAAT CCATGCCAAC CGGTTCCATG TGCTGCCGA GCGGGCATAA ATGCCGTGA  
 5581 CGATCAGCGG TCCAGTGTAC GAAGTTAGGC TGTTAAGAGC CGCGAGCGAT CCTTGAAGCT  
 5641 GTCCCTGATG GTCGTCATCT ACCTGCCCTGG ACAGCATGGC CTGCAACCGC GGCATCCCGA  
 5701 TGCCGCCGGA AGCGAGAAGA ATCATAATGG GGAAGGCCAT CGAGCTCGC GTGCGAACG  
 5761 CCAGCAAGAC GTAGCCCAGC CGCTCGGGCG CCATGCCGGC GATAATGGCC TGCTTCTCGC  
 5821 CGAAACGTTT GGTGGCGGG AAGTTAGGC TGTTAAGAGC CGCGAGCGAT CCTTGAAGCT  
 5881 ATACCGCAAG CGACAGGCCG ATCATCGTC CGCTCCAGCG AAAGCGGTCC TCGCCGAAAA  
 5941 TGACCCAGAG CGCTGCCGGC ACCTGTCCTA CGAGTTGCAT GATAAAGAAG ACAGTCATAA  
 6001 GTGCGGCAC GATAGTCATG CCCCAGGCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC  
 6061 TCAAGGGCAT CGGTCGATCG ACGCTCTCCC TTATGCGACT CCTGCATTAG GAAGCAGCCC  
 6121 AGTAGTAGGT TGAGGCCGTT GAGCACCGCC GCCGCAAGGA ATGGTGCATG CAAGGAGATG-

FIGURE 43C

123/240

6181 GCGCCCAACA GTCCCCCGGC CACGGGGCCT GCCACCATAAC CCACGCCGAA ACAAGCGCTC  
6241 ATGAGCCCGA AGTGGCGAGC CCGA

FIGURE 43D

174/240

pDEST24

## GST carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA

```

1  atc gag atc tcc atc ccc cga aat taa tac gac tca cta tag gga gac cac
tag ctc tag agc tag ggc gct tta [att R1] att atg ctg agt gat atc cgt ctg gtg
52  aac ggt ttc ccc cta gat ccc aag ttt gta cca aaa agc tga acg aga aac
ttg cca aag gga gat cta gtc aaa cat gtt ttg tcc act tgc tct ttg

```

↓

CmR — ccdB — //

att R2

A F L Y K V V I M S

```

1735 // tca ttt tac gtt tct cgt tca gct ttc ttg tac aaa gtt gtt att atg tcc
agt aaa atg cca aqa qca agt cgg aag aac atg ttt cac car taa tac agg
// P I L GST Protein → (~ 223 kDa)
1786 cct ata cta ggt tat tgg aaa att aag ggc ctt gtt cca ccc act cga ctt
gga tat gat cca ata acc ttt taa ttc ccc gaa cac gtt ggg tga gct gaa

```

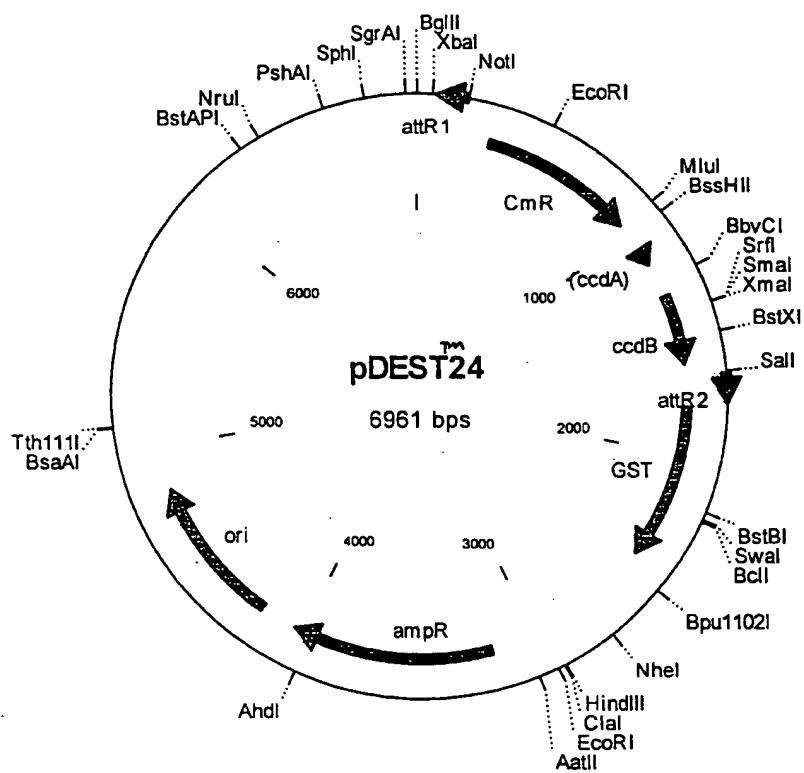


FIGURE 44A

125/240

## pDEST24 6961 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
195..71	attR1
304..963	CmR
1083..1167	inactivated ccdA
1305..1610	ccdB
1651..1775	attR2
1783..2451	GST
3181..4041	ampR
4190..4829	ori

1 ATCGAGATCT CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC  
 61 CCTCTAGATC ACAAGTTTGT ACAAAAAAGC TGAACGAGAA ACGTAAAATG ATATAAATAT  
 121 CAATATATTA AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA  
 181 TATCCAGTCA CTATGGCGGC CGCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCGGC  
 241 TCGTATAATG TGTGGATTTT GAGTTAGGAT CCGGGAGAT TTTCAGGAGC TAAGGAAGCT  
 301 AAAATGGAGA AAAAATCAC TGGATATACC ACCGGTGTATA TATCCCATG GCATCGTAAA  
 361 GAACATTTG AGGCATTCA GTCAAGTGTCAATGTAACCT ATAACCAGAC CGTTCAAGCTG  
 421 GATATTACGG CCTTTTAAAG GACCGTAAAG AAAAATAAGC ACAAGTTTA TCCGGCCTTT  
 481 ATTACATTC TTGCCCCCT GATGAATGCT CATCCGGAT TCCGTATGGC AATGAAAGAC  
 541 GGTGAGCTGG TGATATGGGA TAGTGTTCAC CCGTTTCCA TGAGCAAAC  
 601 GAAACGTTT CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA  
 661 TATTGCGAAG ATGTGGCGTG TTACGGTGAA AACCTGGCCT ATTTCCCTAA AGGGTTTATT  
 721 GAGAATATGT TTTTCTGTC AGCCAATCCC TGGGTGAGTT TCACCAAGTT TGATTTAAC  
 781 GTGGCCAATA TGGACAACCTT CTTGGCCCCC GTTTTCACCA TGGGCAAATA TTATACGCAA  
 841 GGGGACAAGG TGGTGTGTTG GCTGGCATT CAGGGTCATC ATGGCGTCTG TGATGGCTTC  
 901 CATGTCGGCA GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGCG  
 961 TAAACGCGTG GATCCGGCTT ACTAAAAGCC AGATAACAGT ATGCGTATTG GCGCGCTGAT  
 1021 TTTTGGGTA TAAGAATATA TACTGATATG TATACCGAA GTATGTCAAA AAGAGGTGTG  
 1081 CTATGAGCA GCGTATTACA GTGACAGTTG ACAGCGACAG CTATCAGTTG CTCAAGGCAT  
 1141 ATATGATGTC AATATCTCCG GTCTGGTAAG CACAACCATG CAGAATGAAG CCCGTCGTCT  
 1201 GCGTGGCAA CGCTGGAAAG CGGAAAATCA GGAAGGGATG GCTGAGGTG TGAGGTTTAT  
 1261 TGAAATGAAC GCGCTTTTG CTGACGAGAA CAGGGACTGG TGAAATGCAAGTT  
 1321 ACACCTATAA AAGAGAGAGC CGTTATGTC TGTTTGTGGA TGTACAGAGT GATATTATTG  
 1381 ACACGCCCGG GCGACGGATG GTGATCCCCC TGGCCAGTGC ACGCTGCTG TCAGATAAAG  
 1441 TCTCCCGTGA ACTTTACCCG GTGGTGATA TCGGGGATGA AAGCTGGCGC ATGATGACCA  
 1501 CCGATATGGC CAGTGTGCCG GTCTCCGTTA TCGGGGAAAG AGTGGCTGAT CTCAGCCACC  
 1561 GCGAAAATGA CATCAAAAC GCCATTAACC TGATGTTCTG GGGAAATATAA ATGTCAGGCT  
 1621 CCCTTATACA CAGCCAGTCT GCAGGTCGAC CATAGTGAAT GGATATGTTG TGTTTACAG  
 1681 TATTATGTTAG TCTGTTTTTAT ATGCAAATC TAATTAAATA TATTGATATT TATATCATT  
 1741 TACGTTCTC GTTCAGCTTT CTTGTACAAA GTGGTGATTA TGTCCTCTAT ACTAGGTTAT  
 1801 TGGAAAATTA AGGGCTTGTG GCAACCCACT CGACTTCTT TGGAATATCT TGAAAGAAAAA  
 1861 TATGAAGAGC ATTTGTATGA GCGCGATGAA GGTGATAAAT GGCAGAAACAA AAAGTTTGAA  
 1921 TTGGGTTTGG AGTTTCCCAA TCTTCCTTAT TATATTGATG GTGATGTTAA ATTAACACAG  
 1981 TCTATGGCCA TCATACGTTA TATAGCTGAC AAGCACAAACA TGTTGGGTGG TTGTCCAAA  
 2041 GAGCGTGCAG AGATTTCAAT GCTTGAAGGA CGGGTTTGG ATATTAGATA CGGTGTTTCG  
 2101 AGAATTGCAT ATAGTAAAGA CTTTGTAAACT CTCAAAGTTG ATTTTCCTAG CAAGCTACCT  
 2161 GAAATGCTGA AAATGTTCGA AGATCGTTA TGTCATAAAA CATATTAA TGTTGATCAT  
 2221 GTAACCCATC CTGACTTCAT GTTGTATGAC GCTCTGATG TTGTTTTATA CATGGACCCA  
 2281 ATGTCCTGG ATGCGTTCCC AAAATTAGTT TGTTTAAAAA AACGTATTGA AGCTATCCCA  
 2341 CAAATTGATA AGTACTTGAATCCAGCAAG TATATAGCAT GGCCTTTGCA GGGCTGGCAA  
 2401 GCCACGTTTG GTGGTGGCGA CCATCCTCCA AAATCGGATC TGAGTTGGCT GCTGCCACCG CTGAGCAATA  
 2461 TCCGGCTGCT AACAAAGGCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA  
 2521 ACTAGCATAA CCCCTTGGGG CCTCTAAACG GGTTTGAGG GGTTTTTGC TGAAAGGAGG  
 2581 AACTATATCC GGATATCCAC AGGACGGGTG TGGTCGCCAT GATCGCGTAG TCGATAGTGG  
 2641 CTCCAAGTAG CGAAGCGAGC AGGACTGGC GGGGGCCAAA GCGGTCGGAC AGTGCCTCGA-

FIGURE 44B

176/260

2701 GAACGGGTGC GCATAGAAAT TGCATCAACG CATATAGCGC TAGCAGCACG CCATAGTGAC  
 2761 TGGCGATGCT GTCGGAATGG ACGATATCCC GCAAGAGGCC CGGCAGTACC GGCATAACCA  
 2821 AGCCTATGCC TACAGCATCC AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG  
 2881 ATTTCATACA CGGTGCCTGA CTGCGTTAGC AATTAACTG TGATAAACTA CCGCATTAAA  
 2941 GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTGAA GACGAAAGGG CCTCGTGATA  
 3001 CGCCTATTTT TATAGGTTAA TGTATGATA ATAATGGTTT TTAGACGTC AGGTGGCACT  
 3061 TTTCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTT TCTAAATACA TTCAAATATG  
 3121 TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT  
 3181 ATGAGTATTC AACATTCCG TGCGCCCTT ATTCCCTTT TTGCGGCATT TTGCCCTCCT  
 3241 GTTTTGCTC ACCCAGAAC CCGTGGTAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA  
 3301 CGAGTGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC  
 3361 GAAGAACGTT TTCCAATGAT GAGCACTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC  
 3421 CGTGTGACG CGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG  
 3481 GTTGAGTACT CACCAAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA  
 3541 TGCAGTGGCTG CCATAACCAC GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC  
 3601 GGAGGACCGA AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT  
 3661 GATCGTTGGG AACCCGGAGCT GAATGAAGCC ATACCAAACG ACCAGCGTGA CACCACGATG  
 3721 CCTGCAGCAA TGGCAACAAAC GTTGCAGCAA CTATTAACCTG GCGAAGACTACT TACTCTAGCT  
 3781 TCCCAGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAAGGACC ACTTCTGCGC  
 3841 TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAAATCTG GAGCCGGTGA GCGTGGGTCT  
 3901 CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC  
 3961 ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC  
 4021 TCACTGATTA AGCATGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT  
 4081 TAAAAACTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG  
 4141 ACCAAAATCC CTTAACGTGA GTTTTCGTT CACTGAGCGT CAGACCCCGT AGAAAAGATC  
 4201 AAAGGATCTT CTTGAGATCC TTTTTTTCTG CGCGTAACTCT GCTGCTTGCA AACAAAAAAA  
 4261 CCACCGCTAC CAGCGGTGGT TTGTTTGCCTG GATCAAGAGC TACCAACTCT TTTTCCGAAG  
 4321 GTAACTGGCT TCAGCAGAGC GCAGATAACCA AATACTGTCC TTCTAGTGT GCGTAGTTA  
 4381 GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA  
 4441 CCAGTGGCTG CTGCCAGTGG CGATAAGTCC TGTCTTACCG GGTGGACTC AAGACGATAG  
 4501 TTACCCGATA AGGCGCAGCG GTCCGGCTGA ACGGGGGTTT CGTGCACACA GCCCAGCTTG  
 4561 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGT AGCTATGAGA AAGCGCCACG  
 4621 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CGGTAAGCG GCAGGGTCGG AACAGGAGAG  
 4681 CGCACCGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTGT CGGGTTTCGC  
 4741 CACCTCTGAC TTGAGCGTCG ATTTTGTA TGCTCGTCA GGGGGCGGAG CCTATGGAAA  
 4801 AACGCCAGCA ACCGGGCCCTT TTACGGTT CTCGGCTTTTG GCTGGCCTTT TGCTCACATG  
 4861 TTCTTCCCTG CGTTATCCCC TGATTCGTG GATAACCGTA TTACCGCCTT TGAGTGGACT  
 4921 GATACCGCTC GCCGCAGCCG AAACGCCAG CGCAGCGAGT CAGTGAAGCGA GGAAGCGGAA  
 4981 GAGCGCTGA TGCGTAACTT TCTCTTACG CATCTGTGCG GTATTTCAAA CGCGATATAT  
 5041 GGTGCACTCT CAGTACAATC TGCTCTGATG CGCGATAGTT ACGACAGTAT ACACTCCGCT  
 5101 ATCGCTACGT GACTGGGTCA TGGCTGCC CGCACACCCG CCAACACCCG CTGACCGGCC  
 5161 CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG  
 5221 CTGCACTGT CAGAGGTTT CACCGTCATC ACCGAAACGC GCGAGGGAGC TGCGGTAAAG  
 5281 CTCATCAGCG TGCGCTGAA GCGATTCAAA GATGTCTGCC TGTTCATCCG CGTCCAGCTC  
 5341 GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA AAGCGGGCCA TGTTAAGGGC  
 5401 GGTTTTTCTC TGTTGGTCA CTGATGCCCTC CGTGTAAAGGG GGATTTCTGT TCATGGGGGT  
 5461 AATGATAACCG AGAAAACGAG AGAGGATGCT CACGATACGG TTACTGTATG ATGAACATGC  
 5521 CGGTTACTG GAACGTTGTG AGGGTAAACA ACTGGCGGT TGGATGCGGCC GGGACCAAGAG  
 5581 AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAACTACA GATGTAGGTG TTCCACAGGG  
 5641 TAGCCAGCAG CATCCTGCGA TGCAGATCCG GAACATAATG GTGCAGGGCG CTGACTTCCG  
 5701 CGTTTCCAGA CTTTACGAAA CACGGAAACCGA AAGACCAATT CATGTTGTTG CTCAGGTGCC  
 5761 AGACGTTTG CAGCAGCAGT CGCTTCACGT TCGCTCGCGT ATCGGGTATT CATTCTGCTA  
 5821 ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCAAC GACAGGAGCA CGATCATGCG  
 5881 CACCCGTGGC CAGGACCCAA CGCTGCCCGA GATGCGCCGC GTGCGGCTGC TGGAGATGGC  
 5941 GGACGGCATG GATATGTTCT GCCAAGGGTT GGTTTGCCTA TTCACAGTT TCCGCAAGAA  
 6001 TTGATTGGCT CCAATTCTTG GAGTGGTGAA TCCGTTAGCG AGGTGCGGCC GGCTTCATT  
 6061 CAGGTCGAGG TGCCCCGGCT CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT  
 6121 AGGGCGGCCTC CTACAATCCA TGCCAACCG TTCCATGTGC TCGCCGAGGC GGCATAAAATC-

Figure 44C

127/240

6181 GCCGTGACGA TCAGCGGTCC AGTGATCGAA GTIAGGGCTGG TAAGAGCCGC GAGCGATCCT  
6241 TGAAGCTGTC CCTGATGGTC GTCATCTACC TGCCTGGACA GCATGGCTG CAACGCGGCG  
6301 ATCCCGATGC CGCCGGAAGC GAGAAGAAC ATAATGGGGA AGGCCATCCA GCCTCGCGTC  
6361 GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGGGCGCCA TGCCGGCGAT AATGGCTGC  
6421 TTCTCGCCGA AACGTTTGGT GGCGGGACCA GTGACGAAGG CTTGAGCGAG GGCAGTGAAG  
6481 ATTCCGAATA CCGCAAGCGA CAGGCCGATC ATCGTCGCGC TCCAGCGAAA GCGGTCTCG  
6541 CCGAAAATGA CCCAGAGCGC TGCCGGCACCC TGCTTACGA GTTGCATGAT AAAGAAGACA  
6601 GTCATAAGTG CGCGACGAT AGTCATGCC CGCGCCACC GGAAGGAGCT GACTGGTTG  
6661 AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TGCGACTCCT GCATTAGGAA  
6721 GCAGCCCAAGT AGTAGGTTGA GGCGGTTGAG CACCGCCGCC GCAAGGAATG GTGCATGCAA  
6781 GGAGATGGCG CCCAACAGTC CCCCCGGCCAC GGGGCCTGCC ACCATACCCA CGCCGAAACA  
6841 AGCGCTCATG AGCCCGAAGT GGCGAGCCCG ATCTTCCCCA TCGGTGATGT CGGCGATATA  
6901 GGCGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC CGGCAGTAGAG  
6961 G

FIGURE 44D

128/260  
FIGURE 45A

pDEST25

Thioredoxin carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA

1 nag atc tcc atc ccc cga aat tta tac gac cta taa gga gac cac aac  
 ntc tag agc tag ggc gct tta att atg ctg atg qat aca ctt ctg gtg ttg

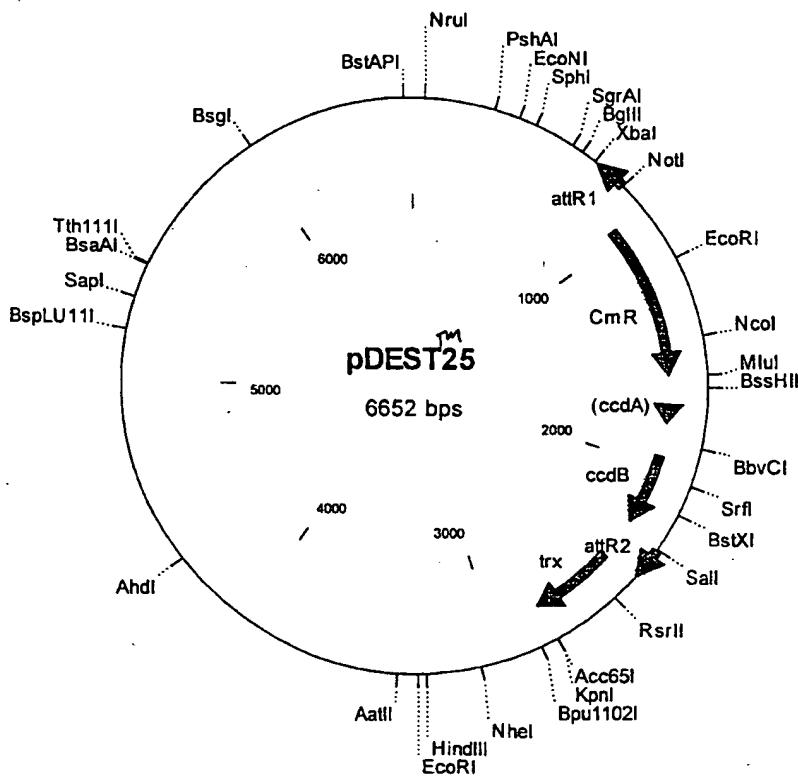
52 ggt ttc cct cta gat cac aag ttt gtt caa aaa agc tga acg aga aac gta  
 cca aag gga gat cta atg ttc aaa cat gtt ttg act tgg tct ttg cat

attR1

1735 // attR2 — A F <sup>\*L</sup> Y K V V I M S D  
 ttt tac gtt tct cgt tca get ttt ttg tac aaa gtg gtt att atg agc gat  
 aaa atg caa aga gca agt cga aag aac atg ttt cac ctc taa tac tcc cta

1786 K I I — Trx Protein (~120 kd) →  
 aaa att att cac ctg act gac gac agt ttt gac acg gat gta ctc aaa gcg  
 ttt taa taa gtg gac tga ctg ctg tca aaa ctg tgc cta cat gag ttt cgc

CmR — ccdB — //



129/240

## pDEST25 6652 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
844..720	attR1
953..1612	CmR
1732..1816	inactivated ccdA
1954..2259	ccdB
2300..2424	attR2
2432..2794	trx

1 CCGGAAGCGA GAAGAACAT AATGGGAAAG GCCATCCAGC CTCGCCTCGC GAACGCCAGC  
 61 AAGACGTAGC CCAGCGCGTC GGCGGCCATG CGGGCGATAA TGGCCTGCTT CTCGCCAAA  
 121 CGTTTGGTGG CGGGACAGT GACGAAGGCT TGAGCGAGGG CGTGAAGAT TCCGAATACC  
 181 GCAAGCGACA GGCGATCAT CGTCGCGCTC CAGCGAAAGC GGTCTCGCC GAAAATGACC  
 241 CAGAGCGCTG CGGGCACCTG TCCTACGAGT TGCAATGATAA AGAACAGACT CATAAGTGCG  
 301 GCGACGATAG TCATGCCCG CGCCCACCGG AAGGAGCTGA CTGGGTTGAA GGCTCTCAAG  
 361 GGCATCGGTC GATCGACGCT CTCCCTTATG CGACTCCTGC ATTAGGAAGC AGCCCAGTAG  
 421 TAGTTGAGG CGTTGAGCA CGCCGCCCGC AAGGAATGGT GCATGCAAGG AGATGGCGCC  
 481 CAACAGTCCC CGGGCACCGG GGCCTGCCAC CATAACCCACG CCGAAACAAAG CGCTCATGAG  
 541 CCCGAAGTGG CGAGCCCGAT CTTCCCCATC GGTGATGTGG GCGATATAGG CGCCAGCAAC  
 601 CGCACCTGTG GCGCCGGTGA TGCCGGCAC GATGCGTCCG GCGTAGAGGAA TCGAGATCTC  
 661 GATCCCGCGA AATTAATACG ACTCACTATA GGGAGACAC AACGGTTTCC CTCTAGATCA  
 721 CAAGTTGTA CAAAAAAAGCT GAACGAGAAA CGTAAATGCA TATAAATATC AATATATTAA  
 781 ATTAGATTTT GCATAAAAAA CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC  
 841 TATGGCGGCC GCATTAGGCA CCCCCAGGCTT TACACTTTAT GCTTCCGGT CGTATAATGT  
 901 GTGGATTGAGG AGTTAGGATC CGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA  
 961 AAAATCACT GGATATACCA CGGTTGATAT ATCCAATGG CATCGTAAAG AACATTTGA  
 1021 GGCATTTCAAG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC  
 1081 CTTTTAAAG ACCGTAAAG AAAATAAGCA CAAGTTTAT CGGGCCTTTA TTCACATTCT  
 1141 TGCCCGCTG ATGAATGCTC ATCCGAATT CGGTATGGCA ATGAAAGACG GTGAGCTGGT  
 1201 GATATGGGAT AGTGTTCACC CTTGTTACAC CGTTTCCAT GAGCAAACGT AAACGTTTC  
 1261 ATCGCTCTGG AGTGAATACC ACGACGATTT CGGCAGTTT CTACACATAT ATTCGCAAGA  
 1321 TGTGGCGTGT TACGGTGAA ACCTGGCTA TTTCCTAAA GGGTTTATG AGAATATGTT  
 1381 TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAAGTTT GATTTAAACG TGGCCAATAT  
 1441 GGACAACCTTC TTCGCCCCCG TTTTCAACAT GGGCAAATAT TATACGCAAG GCGACAAGGT  
 1501 GCTGATGCCG CTGGCGATTC AGGTTCATCA TGCCGCTGT GATGGCTTCC ATGTCCGGCAG  
 1561 AATGTTAAT GAATTACAAAC AGTACTGCGA TGAGTGGCAG GGCGGGGCGT AAACGCGTGG  
 1621 ATCCGGCTTA CTAAAAGCCA GATAACAGTA TGCGTATTG CGCGCTGATT TTTGCGGTAT  
 1681 AAGAATATAT ACTGATATGT ATACCGAAG TATGTCAAA AGAGGTGTGC TATGAAGCAG  
 1741 CGTATTACAG TGACAGTTGA CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA  
 1801 ATATCTCCGG TCTGGTAAGC ACAACCATGC AGAATGAAGC CCGCTGCTG CGTCCGAAC  
 1861 GCTGGAAAGC GGAAAATCAG GAAGGGATGG CTGAGGTCGC CCGGTTTATT GAAATGAACG  
 1921 GCTCTTTGC TGACGAGAAC AGGGACTGGT GAAATGCAGT TTAAGGTTTA CACCTATAAA  
 1981 AGAGAGAGCC GTTATCGTCT GTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCCGGG  
 2041 CGACGGATGG TGATCCCCCT GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCCGTGAA  
 2101 CTTTACCCGG TGGTGCATAT CGGGGATGAA AGCTGGCGA TGATGACCAAC CGATATGGCC  
 2161 AGTGTGCCGG TCTCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC  
 2221 ATCAAAACG CCATTAACCT GATGTTCTGG GGAATATAAA TGTCAGGCTC CCTTATACAC  
 2281 AGCCAGTCTG CAGTCGACCC ATAGTGAACG GATATGTTGT GTTTTACAGT ATTATGTAGT  
 2341 CTGTTTTTA TGCACAT AATTTAATAT ATTGATATT ATATCATTTC ACGTTCTCG  
 2401 TTCAGCTTTTC TTGTCACAAAG TGGTGATTAT GAGCGATAAA ATTATTCAAC TGACTGACGA  
 2461 CAGTTTGTAC ACGGATGTAC TCAAAGCGGA CGGGCGATC CTCGTCGATT TCTGGCAGA  
 2521 GTGGTGCCTG CGGTGCAAAA TGATGCCCG GATTCTGGAT GAAATCGCTG ACAGAATATCA  
 2581 GGGCAAACCTG ACCGTTGCAA AACTGAACAT CGATCAAAC CCTGGCACTG CGCCGAAATA  
 2641 TGGCATCCGT GGTATCCCGA CTCTGCTGCT GTTCAAAAAC GGTGAAGTGG CGGCAACCAA  
 2701 AGTGGGTGCA CTGTCATAAG GTCAGTTGAA AGAGTTCCCTC GACGCTAACCG TGGCCGGTTC  
 2761 TGGTTCTGGT GATGACGATG ACAAGGTACG CGGGGATCGA TCCGGCTGCT AACAAAGCCCC -

Figure 45B

130/240

2821 GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG  
 2881 CCTCTAAACG GGTCTTGAGG GGTTTTTTCG TGAAAGGAGG AACTATATCC GGATATCCAC  
 2941 AGGACGGGTG TGGTCGCCAT GATCGCTAG TCGATAGTGG CTCCAAGTAG CGAACGGCAGC  
 3001 AGGACTGGC GGGGCCAAA GCGGTCGGAC AGTGCCTCGA GAACGGGTGC GCATAGAAAT  
 3061 TGCATCAACG CATATAGCGC TAGCAGCACG CCATAGTGAC TGGCGATGCT GTCCGAATGG  
 3121 ACGATATCCC GCAAGAGGCC CGGCAGTACC GGCATAACCA AGCCTATGCC TACAGCATCC  
 3181 AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG ATTCATACA CGGTGCCTGA  
 3241 CTGCGTTAGC AATTAACTG TGATAAACTA CCGCATTAAA CCTTATCGAT GATAAGCTGT  
 3301 CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA CGCCTATTT TATAGGTTAA  
 3361 TGTATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGAA ATGTGCGCGG  
 3421 AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA  
 3481 ACCCTGATAA ATGCTTCAT AATATTGAAA AAGGAAGAGT ATGAGTATTG AACATTCCG  
 3541 TGTGCCCTT ATTCCCTTT TTGCGGCATT TTGCGCTCCT GTTTTGCTC ACCCAGAAAC  
 3601 GCTGGTAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT  
 3661 GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCAAATGAT  
 3721 GAGCACTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG CGGGGCAAGA  
 3781 GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC  
 3841 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTG TGCACTGCTG CCATAACCAT  
 3901 GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC  
 3961 CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCTT GATCGTTGGG AACCGGAGCT  
 4021 GAATGAAGCC ATACAAACG ACGAGCGTGA CACCAACGATG CCTGCAGCAA TGGCAACAAAC  
 4081 GTTGCGCAA CTATTAACGT GCGAACTACT TACTCTAGCT TCCCGGAAAC AATTAATAGA  
 4141 CTGGATGGAG GCGGATAAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG  
 4201 GTTTATTGCT GATAAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCACT  
 4261 GGGGCCAGAT GTTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGG GTCAGGCAAC  
 4321 TATGGATGAA CGAAATAGAC AGATCGTGA GATAGGTGCC TCACTGATTA AGCATTGGTA  
 4381 ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TAAAAACTTC ATTTTTAATT  
 4441 TAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA  
 4501 GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC  
 4561 TTTTTTCTG CGCGTAATCT GCTGTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT  
 4621 TTGTTGCCG GATCAAGAGC TACCAACTCT TTTCCGAAG GTAACTGGCT TCAGCAGAGC  
 4681 GCAGATACCA AATACTGTCC TTCTAGTGT GCGTAGTTA GGCCACCACT TCAAGAACTC  
 4741 TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG  
 4801 CGATAAGTCG TGTCTTACCG GTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG  
 4861 GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA  
 4921 ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG CTTCCCAGG GGAGAAAGGC  
 4981 GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG  
 5041 GGGAAACGCC TGGTATCTTT ATAGTCTCTG CGGGTTTCGCA CACCTCTGAC TTGAGCGTCG  
 5101 ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGCCCTT  
 5161 TTTACGGTTT CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTCCCTG CGTTATCCCC  
 5221 TGATTCTGTG GATAACCGTA TTACCGCTT TGAGTGGACT GATACCGCTC GCCGCAGCCG  
 5281 AACGACCGAG CGCAGCGAGT CAGTGGCGA GGAAGCGGAA GAGCGCTGA TGCGGTATTT  
 5341 TCTCCTTACG CATCTGTGCG GTATTCACA CCGCATATAT GGTGCACTCT CAGTACAATC  
 5401 TGCTCTGATG CCGCATAGTT AAGCCAGTAT ACACCTCGCT ATCGCTACGT GACTGGTCA  
 5461 TGGCTGCGCC CCGACACCCG CCAACACCCG CTGACCGCGC CTGACGGGCT TGTCTGCTCC  
 5521 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGATGTGT CAGAGTTTT  
 5581 CACCGTCATC ACCGAAACGC GCGAGGGCAGC TGCGGTAAAG CTCATCAGCG TGGTGTGAA  
 5641 GCGATTACA GATGTCTGCG TTGTCATCCG CGTCCAGCTC GTTGAGTTTC TCCAGAACGC  
 5701 TTAATGTCTG GCTTCTGATA AAGCGGGCA TGTTAAGGG GTTTTTTCC TGTTGGTCA  
 5761 CTGATGCTC CCGTGTAAAGGG GGATTTCTGT TCATGGGGT AATGATACCG ATGAAACGAG  
 5821 AGAGGATGCT CACGATACGG GTTACTGATG ATGAACATGC CGGGTTACTG GAACGTTGTG  
 5881 AGGGTAAACAA ACTGGCGGTG TGGATGCGGC GGGACCGAG AAAAATCACT CAGGGTCAAT  
 5941 GCCAGCGCTT CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCTGCGA  
 6001 TGCAGATCCG GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTCCAGA CTTTACGAA  
 6061 CACGGAAACCC GAAGACCATT CATGTTGGT CTCAGGTCGC AGACGTTTG CAGCAGCAGT  
 6121 CGCTTCACGT TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC  
 6181 AGCCTAGCCG GTGCGCTAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA  
 6241 CGCTGCCGA GATGCGCCGC GTGCGGTGCG TGGAGATGGC GGACGCGATG GATATGTTCT-

H6U26 45C

131/240

6301 GCCAAGGGTT GGTTTGCAC TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG  
6361 GAGTGGTGAA TCCGTTAGCG AGGTGCCGCC GGCTTCATT CAGGTCGAGG TGGCCCGGCT  
6421 CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGCAC CTACAATCCA  
6481 TGCCAACCCG TTCCATGTGC TCGCCGAGGC GGCATAAAATC GCCGTGACGA TCAGCGGTCC  
6541 AGTGATCGAA GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC  
6601 GTCATCTACC TGCCTGGACA GCATGGCCTG CAACGGGGC ATCCCGATGC CG

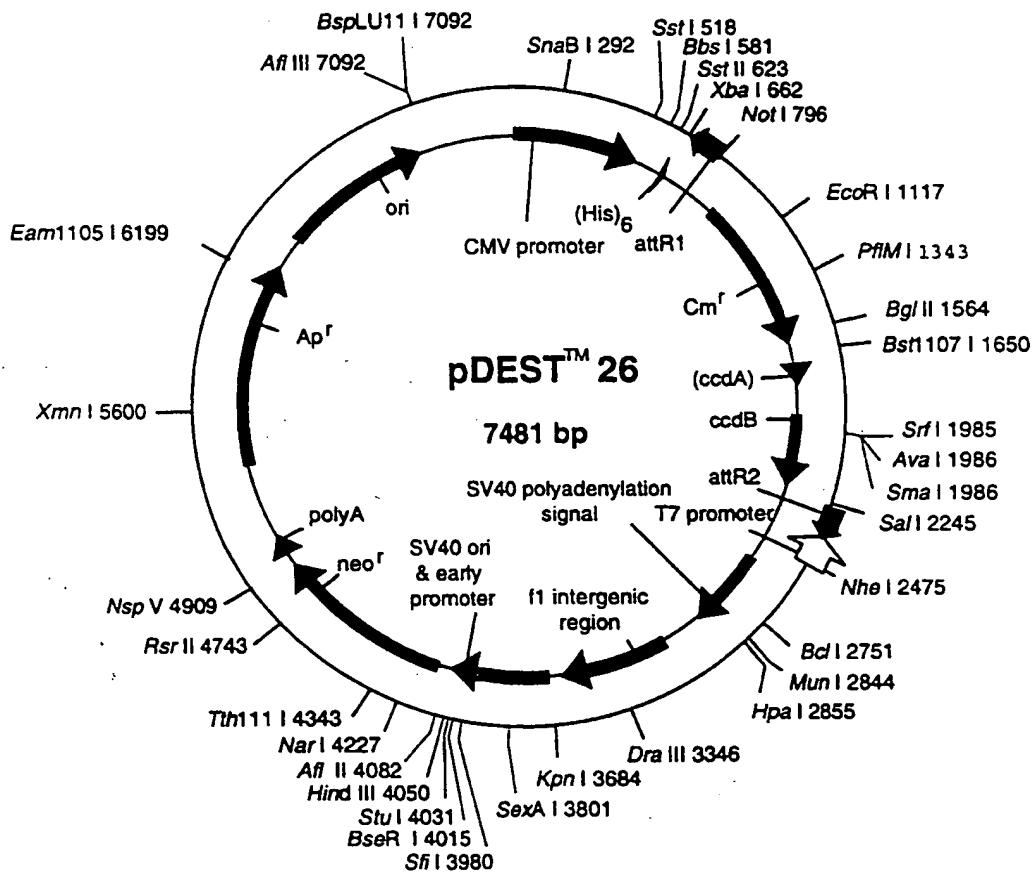
FIGURE 45D

132/240

FIGURE 46A

**pDEST26 His6 Amino Fusion in pCMV Sport-neo Vector**

600 ttg acg tca atg gga gtt tgt ttt ggc aee aaa atc aac ggg act ttc caa  
 aac tgc agt tac cct caa aca aaa ccc tgg ttt tag tgg ccc tga aag gtt  
 651 aat gtc gta aca act ccg ccc cat tga cgc aaa tgg gcg gta ggc gtc tac  
 tta cag cat tgt tga ggg gta act ggg ttt acc cgc cat ccc cac atg  
 702 CMV Promoter mRNA  
 ggt ggg agg tct ata taa gca gag ctc gtt tag tga acc gtc aga tgg ttt  
 cca ccc tcc aga tat att cgt ctc gag caa atc act tgg cgg tct ago gga  
 753 gga gac gcc atc cac gct gtt tgg acc tcc ata gaa gac acc ggg acc gat  
 cct ctg cgg tag gtc cga caa aac tgg agg tat ctt ctg tgg ccc tgg cta  
 804 Start Transl. M A Y Y H H  
 cca gcc tcc gga ctc tag cct agg ccc egg acc latg gcg tac tac cat cac  
 ggt cgg agg cct gag atc gga tcc ggc gec tgg tac cgc atg atg gta gtg  
 H H H H S R S I S Y K K A 20124//  
 855 cat cac cat cac tct aga tca aca agt ttg tac aaa aaa gct gaa gca gaa  
 gta gtg gta gtg aga tct agt tgt tca aac atg ttt ttt cgt ctt gct ctt //  
 Int



133/240

## pDEST26 7481 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
492..509	his6
619..519	attR1
752..1411	CmR
1531..1615	inactivated ccda
1753..2058	ccdB
2099..2223	attR2
2409..2771	SV40 polyA
2966..3421	f1 intergenic region
3485..3903	SV40 promoter
3948..4742	neo
4806..4854	polyA
5265..6125	Apr
6274..6913	ori
7344..385	CMV promoter

1 GTAAACTGCC CACTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA  
 61 CGTCAATGAC GGTAAATGGC CCGCCTGGCA TTATGCCAG TACATGACCT TATGGGACTT  
 121 TCCTACTTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA TGCGGTTTTG  
 181 GCAGTACATC AATGGGGGTG GATAGCGGTT TGACTCACGG GGATTTCAAA GTCTCCACCC  
 241 CATTGACGTC AATGGGAGTT TGTTTGGCA CAAAAATCAA CGGGACTTTC CAAAATGTCG  
 301 TAACAACCTCC GCCCCATTGA CGCAAATGGG CGGTAGGCCT GTACGGTGGG AGGTCTATAT  
 361 AAGCAGAGCT CGTTTAGTGA ACCGTAGAT CGCTTGGAGA CGCCATCCAC GCTGTTTTGA  
 421 CCTCCATAGA AGACACCGGG ACCGATCCAG CCTCCGGACT CTAGCCTAGG CCGCGGACCA  
 481 TGGCGTACTA CCATCACCAT CACCATCACT CTAGATCAAC AAGTTTGAC AAAAAAGCTG  
 541 AACGAGAAC GTAAAATGAT ATAAAATATCA ATATATTAAA TTAGATTTG CATAAAAAC  
 601 AGACTACATA ATACTGTAAA ACACAAACATA TCCAGTCACT ATGGCGGCCG CATTAGGCAC  
 661 CCCAGGCTTT ACACTTTATG CTTCCGGCTC GTATAATGTG TGGATTTTGA GTTGGATCC  
 721 GGCAGAGATT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC  
 781 CGTTGATATA TCCCATTGGC ATCGTAAAGA ACATTTTGAG GCATTTCACT CAGTTGCTCA  
 841 ATGTACCTAT ACCAGACCG TTCAGCTGGA TATTACGGCC TTTTAAAGA CCGTAAAGAA  
 901 AAATAAGCAC AAGTTTTATC CGGCCTTAT TCACATTCTT GCCCGCCTGA TGAATGCTCA  
 961 TCCGGAATTG CGTATGGCAA TGAAAGACGG TGAGCTGGT ATATGGGATA GTGTTCACCC  
 1021 TTGTTACACC GTTTTCCATG AGCAAACCTGA AACGTTTCA TCGCTCTGGA GTGAATACCA  
 1081 CGACGATTTG CGGCAGTTTC TACACATATA TTGCAAGAT GTGGCGTGT ACGGTGAAAA  
 1141 CCTGGCCTAT TTCCCTAAAG GGTTTATTGA GAATATGTT TTCGTCTCAG CCAATCCCTG  
 1201 GGTGAGTTTC ACCAGTTTG ATTTAACGT GGCAATATG GACAACCTCT TCGCCCCCGT  
 1261 TTTCACCATG GGCAAATATT ATACGCAAGG CGACAAGGTG CTGATGCCGC TGGCGATTCA  
 1321 GGTCATCAT GCCGTCTGTG ATGGCTTCCA TGTCGGCAGA ATGCTTAATG AATTACAACA  
 1381 GTACTGCGAT GAGTGGCAGG GCGGGCGTA AAGATCTGGA TCCGGCTTAC TAAAGGCCAG  
 1441 ATAACAGTAT GCGTATTTGC GCGCTGATT TTGCGGTATA AGAATATATA CTGATATGTA  
 1501 TACCCGAAGT ATGTAAAAA GAGGTGTGCT ATGAAGCAGC GTATTACAGT GACAGTTGAC  
 1561 AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA TATCTCCGGT CTGGTAAGCA  
 1621 CAACCATGCA GAATGAAGGCC CGTCGCTCTGC GTGCCGAACG CTGGAAAGCG GAAAATCAGG  
 1681 AAGGGATGGC TGAGGTCGCC CGGTTTATTG AAATGAACGG CTCTTTTGT GACGAGAACCA  
 1741 GGGACTGGTG AAATGCAGTT TAAGGTTTAC ACCTATAAAA GAGAGGCCG TTATCGTCTG  
 1801 TTTGTGGATG TACAGAGTGA TATTATTGAC ACCCCCCGGGC GACGGATGGT GATCCCCCTG  
 1861 GCCAGTGCAC GTCTGCTGTC AGATAAAAGTC TCCCGTGAAC TTTACCGGT GGTGCATATC  
 1921 GGGGATGAAA GCTGGCGCAT GATGACCAACC GATATGGCCA GTGTGCCGGT CTCCGTTATC  
 1981 GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA TCAAAAACGC CATTAACCTG  
 2041 ATGTTCTGGG GAATATAAAAT GTCAGGCTCC CTTATACACA GCCAGTCTGC AGGTGCAGCCA  
 2101 TAGTGAETGG ATATGTTGTG TTTTACAGTA TTATGTAGTC TGTTTTTAT GCAAATCTA  
 2161 ATTAAATATA TTGATATTAA TATCATTTTA CGTTTCTCGT TCAGCTTCT TGTACAAAGT  
 2221 GGTTGATCGC GTGCATGCCA CGTCATAGCT CTCTCCCTAT AGTGAGTCGT ATTATAAGCT  
 2281 AGGCACTGGC CGTCGTTTTA CAACGTCGTG ACTGGGAAA CTGCTAGCTT GGGATCTTG -

FIGURE 46B

134/240

2341 TGAAGGAACC TTACTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA GAGATTAAA  
 2401 GCTCTAAGGT AAATATAAAA TTTTTAAGTG TATAATGTGT TAAACTAGCT GCATATGCTT  
 2461 GCTGCTTGAG AGTTTTGCTT ACTGAGTATG ATTTATGAAA ATATTATACA CAGGAGCTAG  
 2521 TGATTCTAAT TGTGTTGTGTA TTTTAGATTC ACAGTCCCAA GGCTCATTTG AGGCCCTCA  
 2581 GTCCTCACAG TCTGTTCATG ATCATAATCA GCCATACAC ACCTGTAGAG GTTTTACTTG  
 2641 CTTTAAAAAA CCTCCCACAC CTCCCCCTGA ACCTGAAACA TAAAATGAAT GCAATTGTTG  
 2701 TTGTTAACCT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT  
 2761 TCACAAATAA AGCATTTTT TCAC TGCTGGATC GATCCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC  
 2821 TATCTTATCA TGCTGGATC GATCCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC  
 2881 GGTTTGCAGTA TTGGCTGGCG TAATAGCGAA GAGGCCGCA CCGATCGCCC TTCCCAACAG  
 2941 TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAAG CGCGCGGGT  
 3001 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCCTTC  
 3061 GCTTTCTTCC CTTCCCTTCT CGCCACGTTG GCCGGCTTTC CCCGTCAAGC TCTAAATCGG  
 3121 GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT  
 3181 TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTGACG  
 3241 TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCC  
 3301 ATCTCGGTCT ATTCTTTTGTA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA  
 3361 AATGAGCTGA TTTAACAAAT ATTTAACCGC AATTTTAACA AAATATTAAC GTTTACAATT  
 3421 TCGCTGTAT CGGTATTTTC CCCTTACGCA TCTGTGCGGT ATTTACACCC GCATACGCGG  
 3481 ATCTCGCGAG CACCATGGCC TGAAATAACC TCTGAAAGAG GAACCTGGTT AGGTACCTTC  
 3541 TGAGGCGGAA AGAACCAAGCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC  
 3601 TCCCCAGCAG GCAGAAGTAT GCACAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA  
 3661 AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCACTCAA TTAGTCAGCA  
 3721 ACCATAGTCC CGCCCCCTAAC TCCGCCCATC CCGCCCCCTAA CTCCGCCAG TTCCGCCCAT  
 3781 TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCGGCC  
 3841 TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGAAAAG  
 3901 CTTGATTCTT CTGACACAAAC AGTCTCGAAC TTAAGGCTAG AGCCACCATG ATTGAACAAG  
 3961 ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTGCGC TATGACTGGG  
 4021 CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTCCG GCTGTCAGCG CAGGGCGCC  
 4081 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAG GACGAGGCAG  
 4141 CGCGCTATC GTGGCTGGCC ACGACGGCG TTCCCTGCGC AGCTGTGCTC GACGTTGTCA  
 4201 CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGC GGGGCAGGAT CTCTGTCAT  
 4261 CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCCG CGGCTGCATA  
 4321 CGCTTGATCC GGCTACCTGC CCATTGCCACC ACCAACGCAA ACATCGCATC GAGCGAGCAC  
 4381 GTACTCGGAT GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGC  
 4441 TCGGCCAGC CGAACTGTTC GCCAGGCTCA AGGCAGCGCAT GCCCGACGGC GAGGATCTCG  
 4501 TCGTACCCA TGGCGATGCCG TGCTTGGCGA ATATCATGGT GGAAATGGC CGCTTTCTG  
 4561 GATTATCGA CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA  
 4621 CCCGTATAT TGCTGAAGAG CTTGGCCGCG AATGGGCTGA CCGCTTCCCTC GTGCTTACG  
 4681 GTATCGCCGC TCCCGATTTC CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT  
 4741 GAGGGGACT CTGGGGTTCG AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGATG  
 4801 GCCGCAATAA AATATCTTA TTTTCAATTAC ATCTGTGTGT TGGTTTTTG TGTGAATCGA  
 4861 TAGCGATAAG GATCCGCGTA TGGTCACTC TCAGTACAAT CTGCTCTGAT GCCGCATAGT  
 4921 TAAGCCAGCC CCGACACCCG CCAACACCCG CTGACGCCGC CTGACGGGCT TGTCTGCTCC  
 4981 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCACTGTT CAGAGGTTTT  
 5041 CACCGTCATC ACCGAAACGC GCGAGACGAA AGGGCCTCGT GATACGCCATA TTTTATAGG  
 5101 TTAATGTCAT GATAATAATG GTTTCTTAGA CGTCAGGTGG CACTTTTCCG GGAAATGTGC  
 5161 GCGGAACCCC TATTTGTTTA TTTTCTAAA TACATTCAAA TATGATCCG CTCATGAGAC  
 5221 AATAACCTG ATAAATGCTT CAATAATATT GAAAAGGAA GAGTATGAGT ATTCAACATT  
 5281 TCCGTGTCGC CCTTATTCCC TTTTTGCGG CATTTGCGCT TCCCTGTTT GCTCACCCAG  
 5341 AAACGCTGGT GAAAGTAAAAA GATGCTGAAG ATCAGTTGGG TGACGAGTG GGTTACATCG  
 5401 AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTCG CCCCCGAAGAA CGTTTCCAA  
 5461 TGATGAGCAC TTTTAAAGTT CTGCTATGTC GCGCGGTATT ATCCCGTATT GACGCCGGGC  
 5521 AAGAGCAACT CGGTGCCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG  
 5581 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA  
 5641 CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAAC GATCGGAGGA CCGAAGGGAGC  
 5701 TAACCGCTTT TTTGACAAAC ATGGGGGATC ATGTAACCTG CCTTGATCGT TGGGAACCGG  
 5761 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGGCAA -

Figure 4bC

(35/240)

5821 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA  
5881 TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCCTCT GCGCTCGGCC CTTCCGGCTG  
5941 GCTGGTTAT TGCTGATAAA TCTGGAGCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG  
6001 CACTGGGCC AGATGGTAAG CCCTCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG  
6061 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT  
6121 GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAAA CTTCATTTT  
6181 AATTTAAAAG GATCTAGGTG AAGATCCTT TTGATAATCT CATGACCAAAT ATCCCTTAAC  
6241 GTGAGTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG  
6301 ATCCTTTTT TCTGCGCGTA ATCTGCTGCT TGCAACACAA AAAACACCG CTACCAAGCGG  
6361 TGGTTTGTGTT GCCGGATCAA GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA  
6421 GAGCGCAGAT ACCAAATACT GTCCCTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAGA  
6481 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAAGTG GCTGCTGCCA  
6541 GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGC  
6601 AGCGGTGCGGG CTGAACGGGG GGTCGTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA  
6661 CGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA  
6721 AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCAGGAAACAGG AGAGCGCACG AGGGAGCTTC  
6781 CAGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACTTGAGC  
6841 GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACCG  
6901 CCTTTTACG GTTCCTGGCC TTTGCTGGC CTTTGCTCA CATGTTCTTT CCTGCGTTAT  
6961 CCCCTGATTG TGTGGATAAC CGTATTACCG CCTTTGAGTG AGCTGATAACC GCTCGCCGCA  
7021 GCGGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA  
7081 AACCGCCTCT CCCCGCGCGT TGGCCGATTC ATTAATGCGAG AGCTTGCAAT TCGCGCGTTT  
7141 TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGCTCTCAT GAGCGGATAAC ATATTTGAAT  
7201 GTATTAGAA AAATAAACAA ATAGGGGTTTC CGCCACATT TCCCCGAAAAA GTGCCACCTG  
7261 ACGTCTAAGA AACCAATTATT ATCATGACAT TAACCTATAA AAATAGGCAGT AGTACGAGGC  
7321 CCTTTCACTC ATTAGATGCA TGTCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA  
7381 CCGCCCAACG ACCCCCCGCC ATTGACGCTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA  
7441 ATAGGGACTT TCCATTGACCG TCAATGGGTG GAGTATTAC G

FIGURE 46D

136/240  
FIGURE 47A

### pDEST 27 GST Amino Fusion in pCMV Sport-neo Vector

CMV Promoter → mRNA Start

600 nac ggt ggg agg tct ata taa gca gag ctc get tag tga acc gtc aga tcc  
 ntg cca ccc tcc aga tat att cgt ctc gag caa atc act tgg dag tct agc

651 cct gga gac gcc atc cac get gtt ttg acc tcc ata gaa gac acc ggg acc  
 gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt ctg tgg ccc tgg → M A P I L

702 gat cca gcc tcc gga ctc tag cct agg cgg cgg acc atg gcc cct ata cta  
 cta ggt cgg agg cct gag atc gga tcc ggc gcc tgg tgg cgg gga tat gat → Start Transl GST

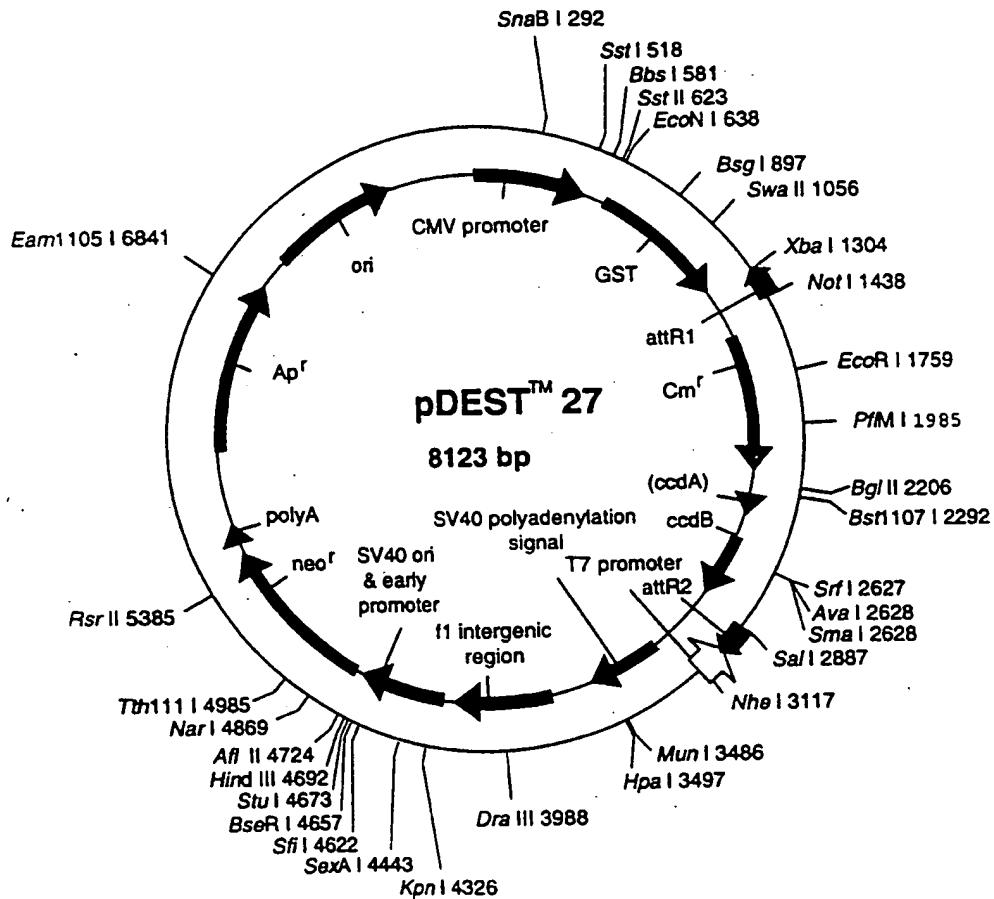
753 ggt tat tgg waa att aag ggc ctt gtg caa ccc act cga ctt ctt ttg gaa  
 cca ata acc ttt taa ttc cgg gaa cac gtt ggg tga gct gaa gaa aac ctt

804 tat ctt gaa gaa aaa tat gaa gag cat ttg tat gag cgc gat gaa ggt gat → H  
 ata gaa ctt ctt ttt ata ctt ctc gta aac ata ctc cgg cta ctt cca cta

1365 ttt ggt ggt ggc gac cat cct cca aaa tcc gat ctg gtt ccc cgt ctc aga  
 aaa cca cca cog ctg gta gga ggt ttt agc cta gac caa ggo gca aga tct → V P R S R

1416 tca aca agt ttg tac aaa aaa gct gaa cga gaa acg  
 agt tgt tca aac atg ttg ttt cga ctt gct ctt tgc → S T S L Y K K A

attR1



137/240

## pDEST27 8123 bp (rotated to position 7800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
130..793	GST
803..927	attR1
1036..1695	CmR
1815..1899	inactivated ccdA
2037..2342	ccdB
2383..2507	attR2
2693..3055	SV40 polyA
3250..3705	f1 intergenic region
3769..4187	SV40 promoter
4232..5026	neo
5090..5138	polyA
5549..6409	Apr
6558..7197	ori
7628..27	CMV promoter

1 ATAAGCAGAG CTCGTTAGT GAACCGTCAG ATCGCCTGGA GACGCCATCC ACGCTGTTTT  
 61 GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGGA CTCTAGCCTA GGCGCGGAC  
 121 CATGGCCCCCT ATACTAGTTT ATTGGAAAAT TAAGGGCTT GTGCAACCCA CTCGACTTCT  
 181 TTTGGAATAT CTTGAAGAAA AATATGAAGA GCATTGTAT GAGCGCGATG AAGGTGATAA  
 241 ATGGCGAAC AAAAGTTTG AATTGGGTTT GGAGTTTCCC AATCTCCTT ATTATATTGA  
 301 TGGTGATGTT AAATTAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA  
 361 CATGTTGGGT GGTTGTCCAA AAGAGCGTGC AGAGATTCA ATGCTTGAAG GAGCGGTTTT  
 421 GGATATTAGA TACGGTGTGTT CGAGAATTGC ATATAGTAAA GACTTTGAAA CTCTAAAGT  
 481 TGATTTCTT AGCAAGCTAC CTGAAATGCT GAAAATGTT GAAAGATCGTT TATGTCATAA  
 541 AACATATTAA AATGGTGATC ATGTAACCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA  
 601 TGTGTTTTA TACATGGACC CAATGTGCCG GGATGCGTTT CCAAAATTAG TTTGTTTTAA  
 661 AAAACGTATT GAAGCTATCC CACAAATTGA TAAGTACTTG AAATCCAGCA AGTATATAGC  
 721 ATGCCCTTG CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCCTC CAAAATCGGA  
 781 TCTGGTCCCG CGTTCTAGAT CAACAAGTTT GTACAAAAAA GCTGAACGAG AACGTAAAA  
 841 TGATATAAAT ATCAATATAT TAAATTAGAT TTGCAATAAA AAACAGACTA CATAATACTG  
 901 TAAAACACAA CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCAGG CTTTACACTT  
 961 TATGCTTCCG GCTCGTAA TGTGTTGATT TTGAGTTAGG ATCCGGCGAG ATTTTCAGGA  
 1021 GCTAAGGAAG CTAAAATGGA GAAAAAAATC ACTGGATATA CCACCGTTGA TATATCCCAA  
 1081 TGGCATCGTA AAGAACATT TGAGGCATT CAGTCAGTT CTCAATGTAC CTATAACCAG  
 1141 ACCGTTCAGC TGGATATTAC GGCTTTTTA AGACCGTAA AGAAAATAA GCACAAGTTT  
 1201 TATCCGGCCT TTATTCACAT TCTTGCCTGC CTGATGAATG CTCATCCGGA ATTCGTATG  
 1261 GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTGTTA CACCGTTTTC  
 1321 CATGAGCAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG  
 1381 TTTCTACACA TATATTCGCA AGATGTGGCG TGTTACGGTG AAAACCTGGC CTATTTCCCT  
 1441 AAAGGGTTA TTGAGAATAT GTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTCACCAGT  
 1501 TTTGATTTAA ACGTGGCCAA TATGGACAAC TTCTTCGCC CCGTTTTCAC CATGGGCAAA  
 1561 TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTTCAC TCATGCCGTC  
 1621 TGTGATGGCT TCCATGTCGG CAGAATGCTT ATGAATTAC AACAGTACTG CGATGAGTGG  
 1681 CAGGGCGGGGG CGTAAAGATC TGGATCCGGC TTACTAAAAG CCAGATAACA GTATGCGTAT  
 1741 TTGCGCGCTG ATTTTGGCG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA  
 1801 AAAAGAGGTG TGCTATGAAG CAGCGTATTAA CAGTGACAGT TGACAGCGAC AGCTATCAGT  
 1861 TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAATGA  
 1921 AGCCCGTCGT CTGCGTGGCG AACGCTGGAA AGCGGAAAT CAGGAAGGGA TGGCTGAGGT  
 1981 CGCCCGGTTT ATTGAAATGA ACGGCTCTT TGCTGACGAG AACAGGGACT GGTGAAATGC  
 2041 AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCGTTATCG TCTGTTGTG GATGTACAGA  
 2101 GTGATATTAT TGACACGCC GGGCGACGGA TGGTGATCCC CCTGGCCAGT GCACGCTCTG  
 2161 TGTCAGATAA AGTCTCCGT GAACCTTACCG CCGTGGTGCA TATCGGGAT GAAAGCTGGC  
 2221 GCATGATGAC CACCGATATG GCCAGTGTGC CCGTCTCCGT TATCGGGAA GAAGTGGCTG  
 2281 ATCTCAGCCA CCGCGAAAAT GACATAAAAA ACGCCATTAA CCTGATGTTC TGGGAAATAT—

Figure 47B

138/240

2341 AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTG ACCATAGTGA CTGGATATGT  
 2401 TGTGTTTAC AGTATTATGT AGTCTGTTT TTATGAAAAA TCTAATTAA TATATTGATA  
 2461 TTTATATCAT TTTACGTTTC TCAGTCAGCT TTCTTGACA AAGTGGTGA TCGCGTCAT  
 2521 GCGACGTCAT AGCTCTCTCC CTATAGTGAG TCGTATTATA AGCTAGGCAC TGGCCGTCGT  
 2581 TTTACAAACGT CGTGACTGGG AAAACTGCTA GCTTGGGATC TTTGTGAAGG AACCTTACTT  
 2641 CTGTGGTGTG ACATAATTGG ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT  
 2701 AAAATTTTTA AGTGTATAAT GTGTTAAACT AGCTGCATAT GCTTGCTGCT TGAGAGTTTT  
 2761 GCTTACTGAG TATGATTATG GAAAATATTA TACACAGGAG CTAGTGATTTC TAATTGTTTG  
 2821 TGTATTAG ATTACAGTC CCAAGGCTCA TTCAGGCCCTC CTAGTCCTC ACAGTCTGTT  
 2881 CATGATCATA ATCAGCCATA CCACATTGT AGAGGTTTA CTTGTTAA AAAACCTCCC  
 2941 ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT GTTGGTTA ACTTGGTTAT  
 3001 TGCAGCTTAT AATGGTTACA AATAAACAA TAGCATCACA AATTCACAA ATAAAGCATT  
 3061 TTTTCACTG CATTCTAGTT GTGGTTGTC CAAACTCATC AATGTATCTT ATCATGTCG  
 3121 GATCGATCCT GCATTAATGA ATCGGCCAAC GCGGGGGAG AGGCGGTTG CGTATTGGCT  
 3181 GGCATAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG  
 3241 GCGAATGGGA CGCGCCCTGT AGCGCGCAT TAAGCGCGC GGGTGTGGTG GTTACCGC  
 3301 GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTC  
 3361 TTCTCGCCAC GTTCGCGGC TTTCCCCGTC AAGCTCTAA TCGGGGGCTC CCTTTAGGGT  
 3421 TCCGATTAG TGCTTACGG CACCTCGACC CAAAAAAACT TGATTAGGGT GATGGTTAC  
 3481 GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTGCCCCTTT GACGTTGGAG TCCACGTTCT  
 3541 TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT  
 3601 TTGATTATA AGGGATTTTG CCGATTCGG CCTATTGGTT AAAAAATGAG CTGATTTAAC  
 3661 AAATATTAA CGCGAATTTT AACAAATAT TAACGTTTAC AATTCGCCT GATGGGTAT  
 3721 TTTCTCCCTA CGCATCTGT CGGTATTCA CACCGCATAC GCGGATCTGC GCAGCACC  
 3781 GGCCTGAAAT AACCTCTGAA AGAGGAACCTT GGTTAGGTAC CTTCTGAGGC GGAAAGAAC  
 3841 AGCTGTGGAA TGTGTGTAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA  
 3901 GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC CCAGGCTCCC  
 3961 CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC  
 4021 TAACTCCGCC CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCG CCCCATGGCT  
 4081 GACTAATTAA TTTTATTAA GCAGAGGCC AGGCCGCCTC GGCCTCTGAG CTATTCCAGA  
 4141 AGTAGTGAGG AGGCTTTTTT GGAGGCTTAG GCTTTGCAA AAAGCTTGAT TCTTCTGACA  
 4201 CAACAGTCTC GAACTTAAGG CTAGAGCCAC CATGATTGAA CAAGATGGAT TGCACCGAGG  
 4261 TTCTCCGGCC GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG  
 4321 CTGCTCTGAT GCCGCCGTGT TCCGGCTGTC AGGCAGGGG CGCCCGGTTT TTTTGTCAA  
 4381 GACCCACCTG TCCGGTGCCTC TGAATGAACT GCAGGACGAG GCAGCGCGC TATCGTGGCT  
 4441 GGCCACGACG GGCGTTCCCTT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAAGGA  
 4501 CTGGCTGCTA TTGGCGAAG TGCCGGGCA GGATCTCTG TCATCTCACC TTGCTCCTGC  
 4561 CGAGAAAGTA TCCATCATGG CTGATGCAAT GCGCGGCTG CATACTGTTG ATCCGGCTAC  
 4621 CTGCCATTG GACCAACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC  
 4681 CGGTCTTGTC GATCAGGATG ATCTGGACGA AGAGCATCG GGGCTCGCGC CAGCCGA  
 4741 GTTCGCCAGG CTCAAGGCCG GCATGCCGA CGCGAGGAT CTGCTGTA CCCATGGCGA  
 4801 TGCCTGCTTG CGAATATCA TGGTGGAAAAA TGGCCGCTTT TCTGGATTCA TCGACTGTGG  
 4861 CCGGCTGGGT GTGGCGGACC GCTATCAGGA CATAGCGTT GCTACCCGTG ATATTGCTGA  
 4921 AGAGCTTGCG GCGAATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCGA  
 4981 TTGCGAGCGC ATCGCCCTCT ATCGCCCTCT TGACGAGTTC TTCTGAGCGG GACTCTGGGG  
 5041 TTGAAATGAA CGGACCAAG GACGCCAAC CTGCCATCAC GATGGCCGA ATAAAATATC  
 5101 TTTATTTCA TTACATCTGT GTGTTGGTTT TTGTTGTGAA TCGATAGCGA TAAGGATCCG  
 5161 CGTATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGA TAGTTAAGCC AGCCCCGACA  
 5221 CCCGCCAACCA CCCGCTGACG CGCCCTGACG GGCTTGCTG CTCCCGGCAT CCGCTTACAG  
 5281 ACAAGCTGTG ACCGCTCCCG GGAGCTGCAT GTGTCAGAGG TTTTACCGT CATCACCGAA  
 5341 ACGCGCGAGA CGAAAGGGCC TCGTGATACG CCTATTAA TAGGTTAAATG TCATGATAAT  
 5401 AATGGTTCT TAGACGTAGT GTGGCAGTT TCAGGGAAAT GTGCGCGGAA CCCCTATTG  
 5461 TTTATTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAAT  
 5521 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTCCGTG TCGCCCTTAT  
 5581 TCCCTTTTGC GCGGCATTTC GCTTCTCTGT TTGCTCAC CGAGAAACGC TGGTGAAAGT  
 5641 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTAACAG  
 5701 CGGTAAGATC CTTGAGAGTT TTGCGCCCGA AGAACGTTT CCAATGATGA GCACTTTAA  
 5761 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTG -

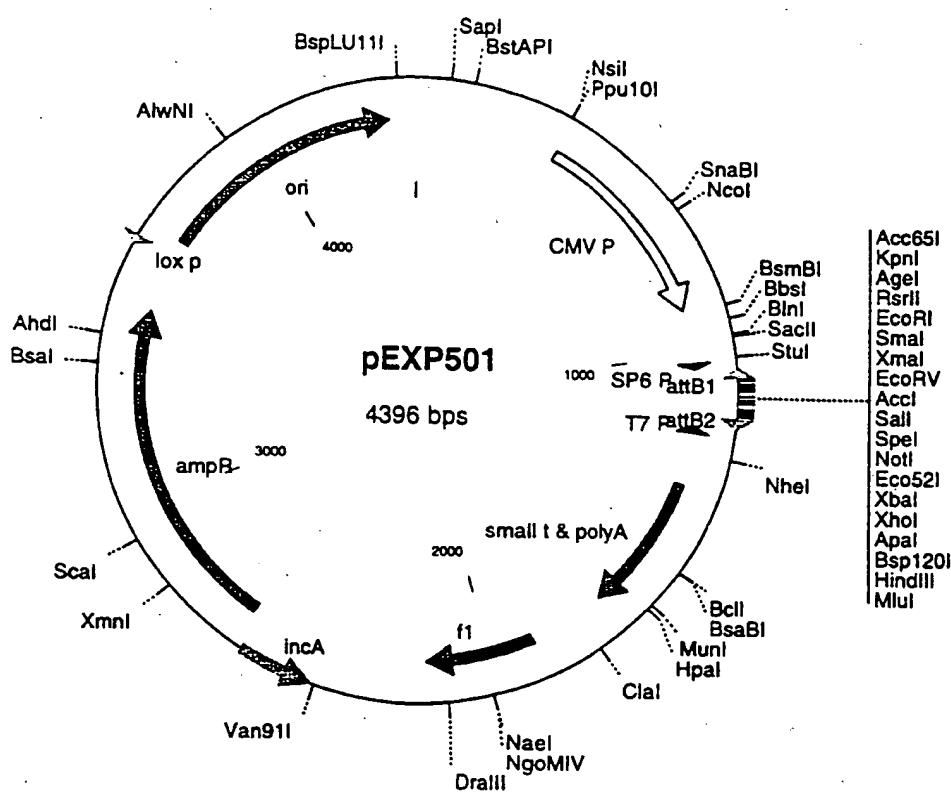
FIGURE 47c

139/240

5821 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT  
 5881 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC  
 5941 TCGGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTITGCA  
 6001 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTGGGAA CCGGAGCTGA ATGAAGCCAT  
 6061 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAAACGT TGCACAAACT  
 6121 ATTAACTGGC GAACACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC  
 6181 GGATAAAGTT GCAGGACCCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA  
 6241 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG  
 6301 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG  
 6361 AAATAGACAG ATCGCTGAGA TAGGTGCGCTC ACTGATTAAG CATTGGTAAC TGTCAACCA  
 6421 AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTAA AAAGGATCTA  
 6481 GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA  
 6541 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTCT TGAGATCCTT TTTTTCTGCG  
 6601 CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCACCA  
 6661 TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATAACAAA  
 6721 TACTGCTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AGAAACTCTG TAGCACCGCC  
 6781 TACATACCTC GCTCTGCTAA TCCGTGTACCG AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG  
 6841 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC  
 6901 GGGGGGTTCG TGACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC TGAGATACCT  
 6961 ACAGCGTGGAG CATTGAGAAA GCGCCACGCT TCCCCAAGGG AGAAAGGCAGG ACAGGTATCC  
 7021 GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCCTG  
 7081 GTATCTTAT AGTCTGTGCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG  
 7141 CTCGTCAGGG GGGGGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCCTTT TACGGTTCCCT  
 7201 GGCCTTTGCG TGGCTTTTG CTCACATGTT CTTTCTGCG TTATCCCCCTG ATTCTGTGGA  
 7261 TAACCGTATT ACCGCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCCAGCG  
 7321 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCCAATA CGCAAACCGC CTCTCCCCGC  
 7381 GCGTTGGCCG ATTCAATTAT GCAGAGCTTGC CAATTGGCGC TTTTTCAAT ATTATTGAAG  
 7441 CATTATCAG GGTATTGTC TCATGAGCGG ATACATATTT GAATGTATT AGAAAAATAA  
 7501 ACAAAATAGGG GTTCCCGCGA CATTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT  
 7561 TATTATCATG ACATTAACCT ATAAATAG GCGTAGTACG AGGCCCTTTC ACTCATTAGA  
 7621 TGCATGTCGT TACATAACTT ACGGTAATG GCCCCCTGG CTGACCGCCC AACGACCCCC  
 7681 GCCCATTGAC GTCAATAATG ACGTATGTT CCATAGTAAC GCGCAATAGGG ACTTTCCATT  
 7741 GACGTCAATG GGTGGAGTAT TTACGGTAACTGCGCTCCTT GGCAGTACAT CAAGTGTATC  
 7801 ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG  
 7861 CCCAGTACAT GACCTTATGG GACTTTCTTA CTTGGCAGTA CATCTACGTA TTAGTCATCG  
 7921 CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGACT  
 7981 CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG GAGTTTGTGTTT TGGCACCAAA  
 8041 ATCAACGGGA CTTTCCAAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGCGGTA  
 8101 GGCGTGTACG GTGGGAGGTC TAT

FIGURE 47)

Figure 4B A: pEXP501: pCMV-SPORT 6 host for attB Libraries



161/260

**Figure 4B:** pEXP5D1 (cont'd).

**Features of the att B cloning vector, pEXP5D1.**  
**Bases within hatched area are replaced by cDNA in some LTI cDNA libraries.**

8v8

---aga gct cgt tta gtg aac cgt cag atc gcc tgg aga cgc cat cca  
 ---tct cga gca aat cac ttc gca gtc tag cgg acc tct gcg gta ggt

→ CMV mRNA

cgc tgt ttt gac ctc cat aga aga cac cgg gac cga tcc agc ctc  
 gcg aca aaa ctg gag gta tct tct gtg gcc ctg get agg tcg gag

Stu I LTI rev primer

cgg act cta gcc tag gcc ggt gag cgg ata aca att tca cac agg  
 gcc tga gat cgg atc cgg cg ctc gcc tat tgt taa agt gtg tcc

ABII rev primer

Stu SP6 promoter → SP6  
 aaa cag cta tga cca tta ggc cta ttt agg tga cac tat aga aca  
 ttt gtc gat act ggt aat cgt gat aaa tcc act gtg ata tct tgt

Int

dK8.1

Age Kpn Rsr II EcoRI Sma I  
 agt ttg tac aaa aaa gca ggc tgg tac cgt tcc gga att ccc ggg  
 tca aac agg ttt ttt cgt cgg act atg gcc agg cct taa ggg ccc

EcoRI Sma I

Spe

Not

Xba

atg /ccg /tgg /agg /agg /tca /gtc /ggc /ggc /ggc /ccc /ttt /agg /ccc /ttt /cat /agg  
 tat /agg /agg /tgc /tgt /gat /cgt /ccg /ccg /ggg /aaa /ttt /tgt /cat /agg

Xba

Apa I

Bpu II

Mlu

dK8.2

Int

ctc gag ggg ccc aag ctt acg cgt acc eag ctt tct tgt aca aag  
 gag ccc ccc ggg ttc gaa tgc gda tgg gtc gaa aga aca tgt ttc

Eco

Pst I

T7

T7 promoter

ABII fwd

Nhe

1272

ttt tac aac gtc gtg act ggg aaa act gtc agc ttg gga tct ttg---  
 aaa atg tgg cag cac tga ccc ttt tga cga tgg aac cct aga aac---

LTI fwd

142/240

## pEXP501 4396 bp

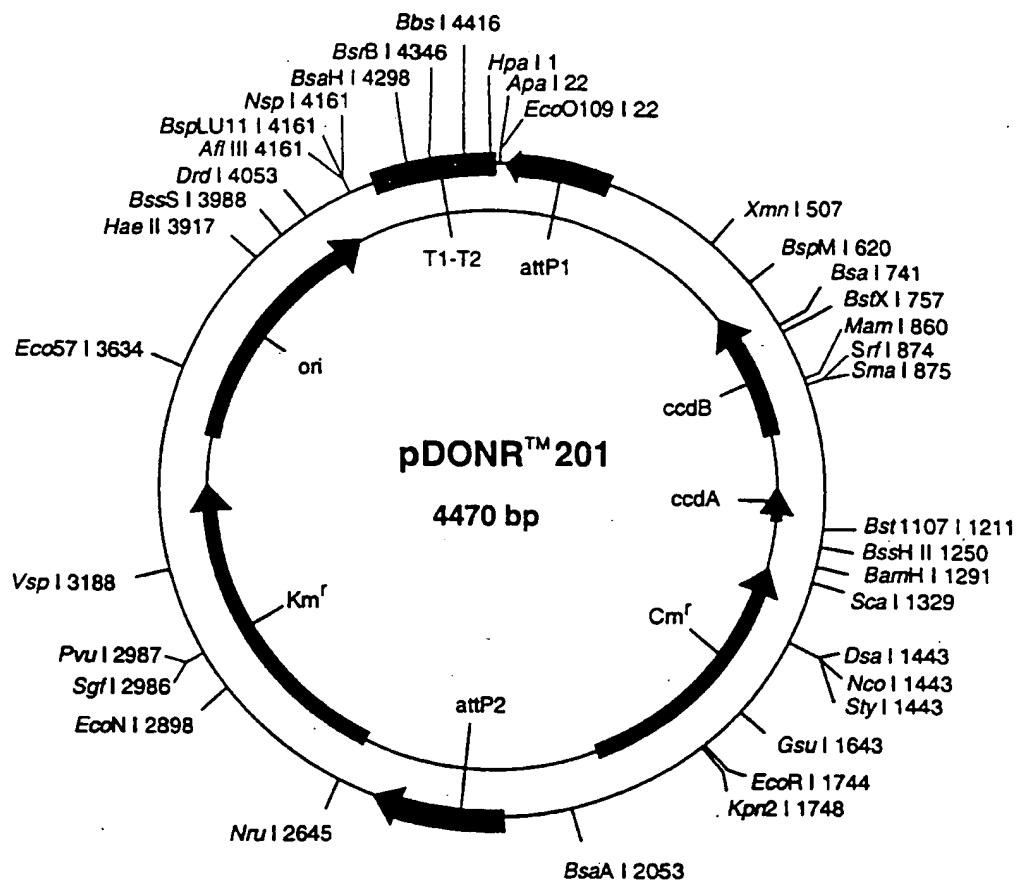
1 CCATTCGCCA TTCAAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT  
 61 ATTACGCCAG CCAATACGCA AACCGCCTCT CCCCGCGCGT TGGCCGATTG ATTAATGCG  
 121 GATCGATCCA GACATGATAA GATACATTGA TGAGTTTGGG CAAACCACAA CTAGAACGCA  
 181 GTGAAAAAAA TGCTTTATTT GTGAAATTG TGATGCTATT GCTTTATTTG TAACCATTAT  
 241 AAGCTGCAAT AAACAAGTTA ACAACAACAA TTGCATTCA TTTATGTTTC AGGTTCAAGG  
 301 GGAGGTGTGG GAGGTTTTTAAAGCAAGTA AAACCTCTAC AAATGTGGTA TGGCTGATTA  
 361 TGATCATGAA CAGACTGTGA GGACTGAGGG GCCTGAAATG AGCCTTGGG CTGTGAATCT  
 421 AAAATACACA AAACAATTAGA ATCACTAGCT CCTGTGTATA ATATTTTCAAAATCATACT  
 481 CAGTAAGCAA AACTCTCAAG CAGCAAGCAT ATGCAGCTAG TTTAACACAT TATACACTTA  
 541 AAAATTTAT ATTTACCTTA GAGCTTTAAA TCTCTGTAGG TAGTTTGTCC AATTATGTCA  
 601 CACCACAGAA GTAAGGTTC TTACACAAAGA TCCCAAGCTA GCAGTTTTC CAGTCACGAC  
 661 GTTGTAACAC GACGGCCAGT GCCTAGCTTA TAATACGACT CACTATAGGG ACCACTTTGT  
 721 ACAAGAAAGC TGGGTACGCG TAAGCTTGGG CCCCTCGAGG GATCCTCTAG AGCGGCCGCG  
 781 GACTAGTGG CTCGTCGACG ATATCCCCGGG AATTCCGGAC CGGTACCAGC CTGCTTTTT  
 841 GTACAAACCTT GTTCTATAGT GTCACCTAAA TAGGCTTAAT GGTCTAGCT GTTTCTGTG  
 901 TGAAATTGTT ATCCGCTCCG CGGCCTAGGC TAGAGTCCGG AGGCTGGATC GGTCCCCTG  
 961 TCTTCTATGG AGGTCAAAAC AGCGTGGATG GCGTCTCCAG GCGATCTGAC GGTTCACTAA  
 1021 ACGAGCTCTG CTTATATAGA CCTCCCCACCG TACACGCCA CCGCCCAATTG GCGTCAATGG  
 1081 GGCGGAGTTG TTACGACATT TTGGAAAGTC CCGTTGATTT TGGTGCACAA ACAAACTCCC  
 1141 ATTGACGTCA ATGGGGTGGG GACTTGGAAA TCCCCGTGAG TCAAACCGCT ATCCACGCC  
 1201 ATTGATGTAC TGCCAAAACC GCATCACCAT GGTAAATAGCG ATGACTAATA CGTAGATGTA  
 1261 CTGCCAAGTA GGAAAGTCCC ATAAGGTCA GTACTGGCA TAATGCCAGG CGGGCCATT  
 1321 ACCGTCATTG ACGTCAATAG GGGCGTACT TGGCATATGA TACACTTGAT GTACTGCCAA  
 1381 GTGGGCAGTT TACCGTAAAT ACTCCACCA TTGACGTCAA TGGAAAGTC CTATTGGCGT  
 1441 TACTATGGGA ACATACGTCA TTATTGACGT CAAATGGCGG GGGTCGTTGG GCGGTCAAGCC  
 1501 AGGCAGGCCA TTACCGTAA GTTATGTAAC GACATGCATC TAATGAGTGA AAGGGCCTCG  
 1561 TACTACGCCCT ATTATTTATAG GTTAATGTC TGATAATAAT GTTTCTTAG ACGTCAGGTG  
 1621 GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTGTTT ATTATTCCTAA ATACATTCAA  
 1681 ATATGTATCC GCTCATGAGA CAATAACCT GATAATGCT TCAATAATAAT TGAAAACGC  
 1741 GCGAATTGCA AGCTCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTGCGTA  
 1801 TTGGCGCTC TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGC  
 1861 GAGCGGTATC AGCTCACTCA AAGGGGTAA TACGGTTATC CACAGAATCA GGGGATAACG  
 1921 CAGGAAAGAA CATGTGAGCA AAAGGCCAG AAAAGGCCAG GAACCGTAAA AAGGCCCGT  
 1981 TGCTGGCGTT TTTCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA  
 2041 GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATAACCA GGCCTTCCC CCTGGAAGCT  
 2101 CCCTCGTGCCTC CTCTCCGTGTT CCGACCTGCG CGCTTACCGG ATACCTGTC GCCTTCTCC  
 2161 CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTA  
 2221 TCGTTGCGCTC CAAGCTGGGC TGTGTGACG AACCCCCCGT TCAGCCGAC CGCTGCGCCT  
 2281 TATCCGGTAA CTATCGTCTT GAGTCAACCG CGGTAAAGACA CGACTTATCG CCACTGGCAG  
 2341 CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA  
 2401 AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA  
 2461 AGCCAGTTAC CTTCGGAAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG  
 2521 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG  
 2581 AAGATCCTTT GATCTTTCT ACAGGGGCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAC  
 2641 GGATTTGGT CATGCCATAA CTTCGTATAG CATACTTAT ACGAAGTTAT GGCATGAGAT  
 2701 TATCAAAAG GATCTTCACC TAGATCTTT TAAATTAAAA ATGAAGTTT AAATCAATCT  
 2761 AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA  
 2821 TCTCAGCGAT CTGTCTATTG CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA  
 2881 CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGTGCA AATGATACCG CGAGACCCAC  
 2941 GCTCACCGGC TCCAGATTAA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA  
 3001 GTGGTCTGCA AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG  
 3061 TAAGTAGTTC GCCAGTTAAAT AGTTTGCAGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG  
 3121 TGTCACTGCTC GTCGTTGGT ATGGCTCATC TCAGCTCCGG TTCCCAACGA TCAAGGCGAG-

FIGURE 48C

143/240

3181 TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG  
3241 TCAGAAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC  
3301 TTACTGTCAT GCCATCCGT AAGATGCTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT  
3361 TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTGCC GGCGTCAATA CGGGATAATA  
3421 CCGGCCACA TAGCAGAACT TAAAAAGTGC TCATCATTGG AAAACGTTCT TCAGGGCGAA  
3481 AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA  
3541 ACTGATCTTC AGCATCTTT ACCTTCACCA GCCTTCTGG GTGAGCAAA ACAGGAAGGC  
3601 AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC  
3661 TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGCCAGGG GTGGGCACAC  
3721 ATATTGATA CCAGCGATCC CTACACAGCA CATAATTCAA TGCGACTTCC CTCTATCGCA  
3781 CATCTTAGAC CTTTATTCTC CCTCCAGCAC ACATCGAAC TGCCGAGCAA GCCGTTCTCA  
3841 CCAGTCCAAG ACCTGGCATG AGCGGATACA TATTGAAATG TATTTAGAAA AATAAACAAA  
3901 TAGGGGTTCC GCGCACATT CCCCCAAAAG TGCCACCTGA AATTGTAAAC GTTAATATTT  
3961 TGTAAAATT CGCGTTAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCAAA  
4021 TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGTCCAG  
4081 TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG  
4141 TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCCCTA ATCAAGTTTT TTGGGGTCGA  
4201 GGTGCCGTAA AGCACTAAAT CGGAACCCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG  
4261 GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGCCTAGGG  
4321 CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCCAC ACCCGCCGCG CTTAATGCAC  
4381 CGCTACAGGG CGCGTC

FIGURE 48D



145/240

## pDONR201 4470 bp (rotated to position 3516)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
260..29	attP1
656..961	ccdB
1099..1184	ccdA
1303..1962	CmR
2210..2442	attP2
2565..3374	Kmr
3495..4134	ori

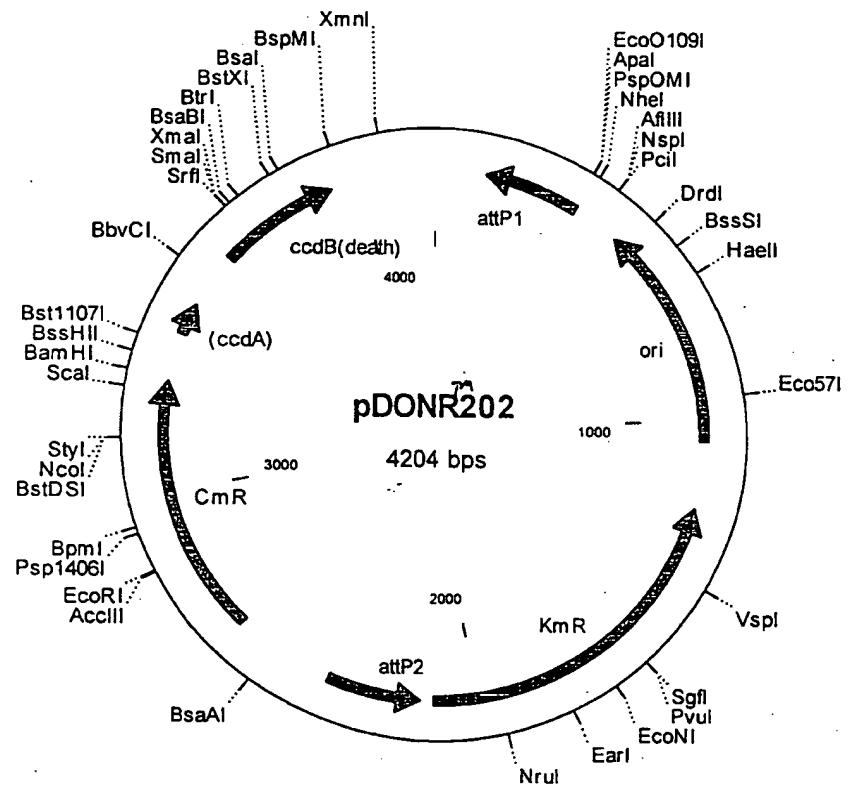
1 GTTAACGCTA GCATGGATCT CGGGCCCCAA ATAATGATT TATTTTGACT GATA GTGACC  
 61 TGTTCGTTGC AACAAATTGA TGAGCAATGC TTTTTTATAA TGCAACTTT GTACAAAAAA  
 121 GCTGAACGAG AAACGTAAAA TGATATAAAT ATCAATATAT TAAATTAGAT TTTGCATAAA  
 181 AAACAGACTA CATAATACTG TAAAACACAA CATATCCAGT CACTATGAAT CAACTACTTA  
 241 GATGGTATTA GTGACCTGTA GTGACCGAC AGCCTTCCAA ATGTTCTTCG GGTGATGCTG  
 301 CCAACTTAGT CGACCGACAG CCTTCCAAAT GTTCTTCTCA AACGGAATCG TC GTATCCAG  
 361 CCTACTCGCT ATTGTCCTCA ATGCCGTATT AAATCATAAA AAGAAATAAG AAAAAGAGGT  
 421 GCGAGCCTCT TTTTTGTGTG ACAAAATAAA AACATCTACC TATT CATATA CGCTAGTGT  
 481 ATAGTCCTGA AAATCATCTG CATCAAGAAC AATTTCACAA CTCTTATACT TTTCTCTTAC  
 541 AAGTCGTTCG GCTTCATCTG GATTTTCAGC CTCTATACTT ACTAAACGTG ATAAAGTTTC  
 601 TGTAATTCT ACTGTATCGA CCTGCAGACT GGCTGTGTAT AAGGGAGCCT GACATTATA  
 661 TTCCCCAGAA CATCAGGTTA ATGGCGTTT TGATGTCATT TCGCGGTGG CTGAGATCAG  
 721 CCACTCTTC CCCGATAACG GAGACGGCA CACTGGCCAT ATCGGTGGTC ATCATGCGCC  
 781 AGCTTCATC CCCGATATGC ACCACGGGGT AAAGTTCACG GGAGACTTTA TCTGACAGCA  
 841 GACGTGCACT GGCCAGGGGG ATCACCATCC GTGCCCGGG CGTGTCAATA ATATCACTCT  
 901 GTACATCCAC AAACAGACGA TAACGGCTCT CTCTTTATA GGTGAAACC TTAAACTGCA  
 961 TTTCACCACT CCTCTGTTCTC GTCAGCAAAA GAGCCGTTCA TTCAATAAA CCGGGCGACC  
 1021 TCAGGCCATCC CTTCTGATT TTCCGCTTTC CAGCGTTCGG CACGCAGACG ACGGGCTTCA  
 1081 TTCTGCATGG TTGTGCTTAC CAGACGGAG ATATTGACAT CATATATGCC TTGAGCAACT  
 1141 GATAGCTGTC GCTGTCAACT GTCACTGTAA TACGCTGCTT CATAGCACAC CTCTTTTG  
 1201 CATACTCGG GTATACATAT CAGTATATAT TCTTATACCG CAAAATCAG CGCGCAAATA  
 1261 CGCATACTGT TATCTGGCTT TTAGTAAGCC GGATCCACGC GATTACGCC CGCCCTGCCA  
 1321 CTCATCGCAG TACTGTTGTA ATTCAATTAG CATTCTGCC ACATGGAAGC CATCACAGAC  
 1381 GGCATGATGA ACCTGAATCG CCAGCGCAT CAGCACCTTG TCGCCTGCG TATAATATT  
 1441 GCCCATGGTG AAAACGGGGG CGAAGAAGTT GTCCATATTG GCCACGTTA AATCAAAACT  
 1501 GGTGAAACTC ACCCAGGGAT TGGCTGAGAC GAAAAACATA TTCTCAATAA ACCCTTTAGG  
 1561 GAAATAGGCC AGGTTTTCAC CGTAACACGC CACATCTTGC GAATATATGT GTAGAAACTG  
 1621 CCGGAAATCG TCGTGGTATT CACTCCAGAG CGATGAAAAC GTTTCAGTTT GTCATGGAA  
 1681 AACGGTGTAA CAAGGGTGA CACTATCCA TATCACCAGC TCACCGTCTT TCATTGCCAT  
 1741 ACGGAATTCC GGATGAGCAT TCATCAGGCG GGCAGAATG TGAATAAAAGG CCGGATAAAA  
 1801 CTTGTGCTTA TTTTTCTTTA CGGTCTTAA AAAGGCCGTA ATATCCAGCT GAACGGTCTG  
 1861 GTTATAGGTA CATTGAGCAA CTGACTGAAA TGCCCTCAAA TGTTCTTAC GATGCCATTG  
 1921 GGATATATCA ACGGTGGTAT ATCCAGTGAT TTTTTCTCC ATTTTAGCTT CCTTAGCTCC  
 1981 TGAAATCTC GATAACTCAA AAAATACGCC CGGTAGTGAT CTTATTTCAT TATGGTGA  
 2041 GTTGGAACCT CTTACGTGCC GATCAACGTC TCATTTTGC CAAAAGTTGG CCCAGGGCTT  
 2101 CCCGGTATCA ACAGGGACAC CAGGATTAT TTATTCAGCG AAGTGATCTT CCGTCACAGG  
 2161 TATTTATTGCG GCGCAAAGTG CGTCGGTGA TGCTGCCAAC TTAGTCGACT ACAGGTCACT  
 2221 AATACCATCT AAGTÀGTTGA TTCATAGTGAT CTGGATATGT TGTGTTTAC AGTATTATGT  
 2281 AGTCTGTTTT TTATGCAAAA TCTAATTAA TATATTGATA TTTATATCAT TTTACGTTTC  
 2341 TCGTTCAGCT TTCTTGACAA AAGTGGCAT TATAAGAAAG CATTGCTTAT CAATTGTTG  
 2401 CAACGAACAG GTCACTATCA GTCAAATAA AATCATTATT TGCCATCCAG CTGCACTCT  
 2461 GGCCCGTGTGTC TCAAAATCTC TGATGTTACA TTGCAACAGA TAAAATATA TCATCATGAA  
 2521 CAATAAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC CATATTCAAC  
 2581 GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT GGGTATAAAT  
 2641 GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGTAT GGGAAAGCCCG  
 2701 ATGCGCCAGA GTTGTCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT GTTACAGATG -

FIGURE 49B

2761 AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTATA  
2821 TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTCGCAT CCCCGGAAAA ACAGCATTCC  
2881 AGGTATTAGA AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG GCAGTGTCC  
2941 TGCGCCGGTT GCATTCGATT CCTGTTGTA ATTGTCCCTT TAACAGCGAT CGCGTATTC  
3001 GTCTCGCTCA GGCGCAATCA CGAACATGAATA ACGGTTGGT TGATGCGAGT GATTTGATG  
3061 ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAAA CTTTGCCAT  
3121 TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTGACG  
3181 AGGGAAATT AATAGGTTGT ATTGATGTTG GACCGAGTCGG AATCGCAGAC CGATACCAGG  
3241 ATCTTGCCAT CCTATGGAAC TGCCCTGGTG AGTTTCTCC TTCATTACAG AAACGGCTT  
3301 TTCAAAAATA TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTCAT TTGATGCTCG  
3361 ATGAGTTTTT CTAATCAGAA TTGGTTAATT GGTTGTAACA CTGGCAGAGC ATTACGCTGA  
3421 CTTGACGGGA CGGCGCAAGC TCATGACCAA AATCCCTAA CGTGAGTTT CGTTCCACTG  
3481 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT  
3541 AATCTGCTGC TTGCAAACAA AAAAACACC GCTACCGCG GTGGTTGGT TGCCGGATCA  
3601 AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC  
3661 TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG AACTCTGTAG CACCGCCTAC  
3721 ATACCTCGCT CTGCTAATCC TGTTACCACT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT  
3781 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTGGT GCTGAACGGG  
3841 GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACTACA  
3901 GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT  
3961 AAGGGCAGG GTCGGAACAG GAGAGGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA  
4021 TCTTTATAGT CCTGTCGGGT TTGCCCCACT CTGACTTGAG CGTCGATTTT TGTGATGCTC  
4081 GTCAGGGGGG CGGAGCCTAT GGAAAACGC CAGCAACCGCG GCCTTTTAC GGTTCTGGC  
4141 CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA  
4201 CCGTATTACC GCTAGCCAGG AAGAGTTGT AGAAACGCAA AAAGGCCATC CGTCAGGATG  
4261 GCCTTCTGCT TAGTTTGATG CCTGGCAGTT TATGGCGGGC GTCCCTGCCG CCACCCCTCCG  
4321 GGCGTTGCT TCACAACGTT CAAATCCGCT CCCGGCGGAT TTGTCCTACT CAGGAGAGCG  
4381 TTCACCGACA ACAACAGAT AAAACGAAAG GCCCAGTCTT CCGACTGAGC CTTTCGTTTT  
4441 ATTTGATGCC TGGCAGTTCC CTACTCTCGC

FIGURE 49C

147/240  
FIGURE 50A: pDONR202 (KanR)



## pDONR202 4204 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
369..127	attP1
486..1059	ori
1228..2107	KmR
2381..2140	attP2
2629..3288	CmR
3408..3492	inactivated ccdA
3630..3935	ccdB

1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTTGAG AAGAACATTT  
 61 GGAAGGCTGT CGGTCGACTA AGTTGGCAGC ATCACCCGAA GAACATTGG AAGGCTGTCG  
 121 GTCGACTACA GGTCACTAAT ACCATCTAAG TAGTTGATTC ATAGTGAATG GATATGTTGT  
 181 GTTTTACAGT ATTATGTAGT CTGTTTTTA TGCAAAATCT AATTTAATAT ATTGATATTT  
 241 ATATCATTTC ACGTTTCTCG TTCAGCTTT TTGTCACAAAG TTGGCATTAT AAAAAGCAT  
 301 TGCTCATCAA TTGTTGCAA CGAACAGGT ACTATCAGTC AAAATAAAAT CATTATTGG  
 361 GGCCCCGAGAT CCATGCTAGC GGTAATACGG TTATCCACAG ATCAGGGGA TAACGCCAGGA  
 421 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG  
 481 GCGTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG  
 541 AGGTGGCAGA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC  
 601 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCCTT TCTCCCTTCG  
 661 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTTAGGTATC TCAGTTCGGT GTAGGTCGTT  
 721 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAAGC CCGACCGCTG CGCCTTATCC  
 781 GGTAACATATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC  
 841 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG  
 901 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTGGTA TCTGCGCTCT GCTGAAGCCA  
 961 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCC CGCTGGTAGC  
 1021 GGTGGTTTTT TTGTTGCAA GCAGCAGATT ACGCGCAGAA AAAAGGAGTC TCAAGAAGAT  
 1081 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCAGC TTAAGGGATT  
 1141 TTGGTCATGA GCTTGCAGCG TCCCGTCAAG TCAGCGTAAT GCTCTGCCAG TGTTACACC  
 1201 AATTAACCAA TTCTGATTAG AAAAACTCAT CGAGCATCAA ATGAAACTGC AATTTATTCA  
 1261 TATCAGGATT ATCAATACCA TATTTTGAA AAAGCCGTT CTGTAATGAA GGAGAAAATC  
 1321 CACCGAGGCA GTTCCATAGG ATGGCAAGAT CCTGGTATCG GTCTGCGATT CGCACTCGTC  
 1381 CAACATCAAT ACAACTATT AATTTCCCT CGTCAAAAAT AAGGTTATCA AGTGAGAAAT  
 1441 CACCATGAGT GACGACTGA TCCGGTGAGA ATGGCAAAAG TTTATGCATT TCTTTCCAGA  
 1501 CTTGTTCAAC AGGCCAGCA TTACGCTCGT CATCAAAATC ACTCGCATCA ACCAAACCGT  
 1561 TATTCAATTG TGATTGCGC TGAGCGAGAC GAAATACGCG ATCGCTGTTA AAAGGACAAT  
 1621 TACAAACAGG AATCGAATG AACCAGGCGCA GGAACACTGC CAGCGCATCA ACAATATTT  
 1681 CACCTGAATC AGGATATTCT TCTAATACCT GGAATGCTGT TTTCCGGGG ATCGCAGTGG  
 1741 TGAGTAACCA TGCATCATCA GGAGTACCGA TAAAATGCTT GATGGTCGGA AGAGGCATAA  
 1801 ATTCCGTCAG CCAGTTAGT CTGACCATCT CATCTGTAAC ATCATTGGCA ACGCTACCTT  
 1861 TGCCATGTTT CAGAAACAAC TCTGGCGCAT CGGGCTTCCC ATACAAGCGA TAGATTGTCG  
 1921 CACCTGATTG CCCGACATTA TCGCGAGCCC ATTATACCC ATATAAAATCA GCATCCATGT  
 1981 TGGAATTAA TCGCGGCCTC GACGTTTCCC GTTGAATATG GCTCATAAACA CCCCTTGTAT  
 2041 TACTGTTTAT GTAAGCAGAC AGTTTATTG TTCATGATGA TATATTTTA TCTTGTGCAA  
 2101 TGTAACATCA GAGATTTGA GACACGGGCC AGAGCTGCAG CTGGATGGCA AATAATGATT  
 2161 TTATTTGAC TGATAGTGAC CTGTTGTTG CAACAAATTG ATAAGCAATG CTTTCTTATA  
 2221 ATGCCAACTT TGACAAGAA AGCTGAACGA GAAACGTAAA ATGATATAAA TATCAATATA  
 2281 TTAAATTAGA TTTGCATAA AAAACAGACT ACATAATACT GTAAAACACA ACATATCCAG  
 2341 TCACTATGAA TCAACTACTT AGATGGTATT AGTGACCTGT AGTCGACTAA GTTGGCAGCA  
 2401 TCACCCGACG CACTTGCAGC CGAATAAATA CCTGTGACGG AAGATCACTT CGCAGAATAA  
 2461 ATAAATCCTG GTGTCCTGT TGATACCGGG AAGCCCTGGG CCAACTTTTG GCGAAAATGA  
 2521 GACGTGATC GGACGTAAAG AGGTCCAAC TTTCACCCATA ATGAAATAAG ATCACTACCG  
 2581 GGCATTTTT TTGAGTTATC GAGATTTCA GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA  
 2641 ATCACTGGAT ATACCACCGT TGATATATCC CAATGGCATC GTAAAGAACAA TTTTGAGGCA  
 2701 TTTCAGTCAG TTGCTCAATG TACCTATAAC CAGACCGTTC AGCTGGATAT TACGGCCTT-

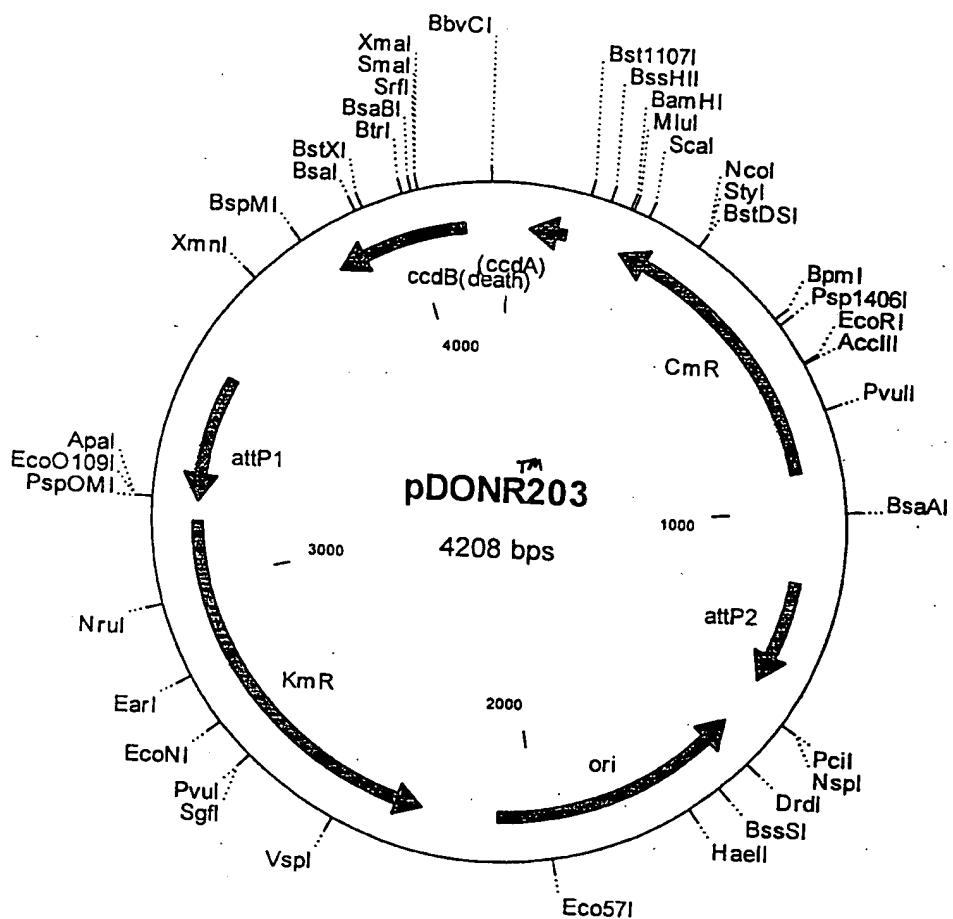
Figure 50B

149/240

2761 TTAAAGACCG TAAAGAAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTGCC  
2821 CGCCTGATGA ATGCTCATCC GGAATTCGT ATGGCAATGA AAGACGGTA GCTGGTGATA  
2881 TGGGATAGTG TTCACCCCTTG TTACACCGTT TTCCATGAGC AAACGTAAAC GTTTTCATCG  
2941 CTCTGGAGTG AATACCACGA CGATTTCCGG CAGTTCTAC ACATATATTG GCAAGATGTG  
3001 GCGTGTACG GTGAAAACCT GGCTTATTTT CCTAAAGGGT TTATTGAGAA TATGTTTTTC  
3061 GTCTCAGCCA ATCCCTGGGT GAGTTTCAAC AGTTTGATT TAAACGTGGC CAATATGGAC  
3121 AACTCTTCG CCCCCGTTT CACCATGGGC AAATATTATA CGCAAGGCGA CAAGGTGCTG  
3181 ATGCCGCTGG CGATTCAAGGT TCATCATGCC GTCTGTGATG GCCTCCATGT CGGCAGAATG  
3241 CTTAATGAAT TACAACAGTA CTGCGATGAG TGGCAGGGCG GGGCGTAATC GCGTGGATCC  
3301 GGCTTACTAA AAGCCAGATA ACAGTATGCG TATTGCGCG CTGATTTTG CGGTATAAGA  
3361 ATATATACTG ATATGTATAC CCGAAGTATG TCAAAAGAG GTGTGCTATG AAGCAGCGTA  
3421 TTACAGTGAC AGTTGACAGC GACAGCTATC AGTTGCTCAA GGCAATATG ATGTCAATAT  
3481 CTCCGGTCTG GTAAGCACAA CCATGCAGAA TGAAGCCCGT CGTCTGCGTG CGAACGCTG  
3541 GAAAGCGGAA AATCAGGAAG GGATGGCTGA GGTCGCCCCG TTTATTGAAA TGAACGGCTC  
3601 TTTTGCTGAC GAGAACAGGG ACTGGTAAA TGCAAGTTAA GGTTTACACC TATAAAAGAG  
3661 AGAGCCGTTA TGTTCTGTT GTGGATGTAC AGAGTGTAT TATTGACACG CCCGGGCGAC  
3721 GGATGGTGAT CCCCCTGGCC AGTGCACGTC TGCTGTAGA TAAAGTCTCC CGTGAACCTT  
3781 ACCCGGTGGT GCATATCGGG GATGAAAGCT GGCGCATGAT GACCACCGAT ATGGCCAGTG  
3841 TGCCGGTCTC CGTTATCGGG GAAGAAGTGG CTGATCTCAG CCACCGCGAA AATGACATCA  
3901 AAAACGCCAT TAACCTGATG TTCTGGGAA TATAAATGTC AGGCTCCCTT ATACACAGCC  
3961 AGTCTGCAGG TCGATAACAGT AGAAATTACA GAAACTTTAT CACGTTTAGT AAGTATAGAG  
4021 GCTGAAAATC CAGATGAAGC CGAACGACTT GTAAGAGAAA AGTATAAGAG TTGTGAAATT  
4081 GTTCTTGATG CAGATGATT TCAGGACTAT GACACTAGCG TATATGAATA GGTAGATGTT  
4141 TTTATTTTGT CACACAAAAA AGAGGCTCGC ACCTTTTTT CTTATTTCTT TTTATGATTT  
4201 AATA

FIGURE 50C

FIGURE 51A pDONR203 (kanR)



151/240

## pDONR203 4208 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
47..131	inactivated ccdA
251..910	CmR
1158..1398	attP2
1509..2082	ori
2251..3130	KmR
3464..3174	attP1
3812..4117	ccdB

1 GCGTTGGCA CGCAGACGAC GGGCTTCATT CTGCATGGTT GTGCTTACCA GACCGGAGAT  
 61 ATTGACATCA TATATGCCCTT GAGCAACTGA TAGCTGTCGC TGTCAACTGT CACTGTAATA  
 121 CGCTGCTTCA TAGCACACCT CTTTTGACA TACTTCGGGT ATACATATCA GTATATATT  
 181 TTATACCGCA AAAATCAGCG CGCAAATACG CATACTGTTA TCTGGCTTT AGTAAGCCGG  
 241 ATCCACCGGT TTACGCCCGG CCCTGCCACT CATCGCAGTA CTGTTGTAAT TCATTAAGCA  
 301 TTCTGCCGAC ATGGAAGCCA TCACAGACGG CATGATGAAC CTGAATGCC AGCGGCATCA  
 361 GCACCTTGTC GCCTTGCGTA TAATATTGTC CCATGGTGA AACGGGGGGG AAGAACGTTG  
 421 CCATATTGGC CACGTTAAA TCAAAACTGG TGAAACTCAC CCAGGGATTG GCTGAGACGA  
 481 AAAACATATT CTCATAAAC CTTTGGGA AATAGGCCAG GTTTTCACCG TAACACGCCA  
 541 CATCTTGCGA ATATATGTGT AGAAACTGCC GGAAATCGTC GTGGTATTCA CTCCAGAGCG  
 601 ATGAAAACGT TTCAGTTGTC TCATGGAAAA CGGTGTAACA AGGGTGAACA CTATCCCATA  
 661 TCACCAAGCTC ACCGTCTTTC ATTGCCATAC GGAATTCCGG ATGAGCATTG ATCAGGCCGG  
 721 CAAGAAATGTG ATAAGGAGCC GGATAAAACT TGTGCTTATT TTTCTTACG GTCTTAAAA  
 781 AGGCCGTAAT ATCCAGCTGA ACGGCTGGT TATAGGTACA TTGAGCAACT GACTGAAATG  
 841 CCTCAAAATG TTCTTACGA TGCCATTGGG ATATATCAAC GGTGGTATAT CCAGTGATTT  
 901 TTTCTCCAT TTAGCTTCCG AAAATCTCGA TAACTCAAAA AATACGCCGG  
 961 GTAGTGATCT TATTCATTA TGGTGAAGT TGGAACCTCT TACGTGCCGA TCAACGTCTC  
 1021 ATTTTCGCCA AAAGTTGGCC CAGGGCTTCC CGGTATCAAC AGGGACACCA GGATTATTT  
 1081 ATTCTGCGAA GTGATCTTCC GTCACAGGTA TTTATCGGC GCAAAGTGGC TCGGGTGATG  
 1141 CTGCCAACTT AGTCGACTAC AGGTCACTAA TACCATCTAA GTAGTTGATT CATACTGACT  
 1201 GGATAATGTTG TGTGTTACAG TATTATGTAG TCTGTTTTT ATGCAAAATC TAATTAAATA  
 1261 TATTGATATT TATATCATT TACGTTCTC GTTCAGCTT CTTGTACAAA GTGGCATT  
 1321 TAAGAAAGCA TTGCTTATCA ATTTGTTGCA ACGAACAGGT CACTATCAGT CAAAATAAAA  
 1381 TCATTATITG CCATCCAGCT AGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA  
 1441 GGAAAGAACCA TGTGAGCAA AGGCCAGCAA AAGGCCAGGA ACCGTAACCGG GGCCGCGTTG  
 1501 CTGGCGTTT TCCATAGGCT CCGCCCCCT GACGAGCATE ACACAAATCG ACGCTCAAGT  
 1561 CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTCCCCC TGGAAAGCTCC  
 1621 CTCGTGCGCT CTCCCTGTC GACCCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT  
 1681 TCGGGAAGCG TGGCGTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTG GGTGTAGGTC  
 1741 GTTCGCTCCA AGCTGGGCTG TGTGACGAA CCCCCCGTT AGCCCGACCG CTGCGCCTTA  
 1801 TCCGTAACT ATCGTCTGCA GTCCAACCCG GTAAGACACG ACTTATGCC ACTGGCAGCA  
 1861 GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTTAGGCG GTGCTACAGA GTTCTGAAG  
 1921 TGGTGGCTA ACTACGGCTA CACTAGAAGA ACAGTATTG GTATCTGCGC TCTGCTGAAG  
 1981 CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTGATCCG GCAAACAAAC CACCGCTGGT  
 2041 AGCGGTGGTT TTTTGTGTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA  
 2101 GATCCTTGA TCTTCTAC GGGGTCTGAC GCTCAGTGGA ACGBAAACTC ACGTTAAGGG  
 2161 ATTTGGTCA TGAGCTTGC CCGTCCCGTC AAGTCAGCGT AATGCTCTGC CAGTGTAC  
 2221 ACCAATTAAC CAATCTGA TAGAAAAACT CATCGAGCAT CAAATGAAAC TGCAATTAT  
 2281 TCATATCAGG ATTATCAAT CCATATTGTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA  
 2341 ACTCACCGAG GCAGTTCCAT AGGATGCCAA GATCCTGGTA TCGGTCTGCG ATTCCGACTC  
 2401 GTCCAACATC AATACAAACCT ATTAATTCC CCTCGTCAA AATAAGGTTA TCAAGTGAGA  
 2461 AATCACCATG AGTGACGACT GAATCCGGTG AGAATGCCAA AAGTTTATGC ATTTCTTCC  
 2521 AGACTTGTTC AACAGGCCAG CCATTACGCT CGTCATCAAA ATCACTGCC TCAACCAAAC  
 2581 CGTTATTCACT CGTGATTGC GCCTGAGCGA GACGAAATAC GCGATCGCTG TTAAAAGGAC  
 2641 AATTACAAAC AGGAATCGAA TGCAACCGGC GCAGGAACAC TGCCAGCGCA TCAACAAAT  
 2701 TTTCACCTGA ATCAGGATAT TCTTCTAATA CCTGGAATGC TGTTTTCCG GGGATCGCAG-

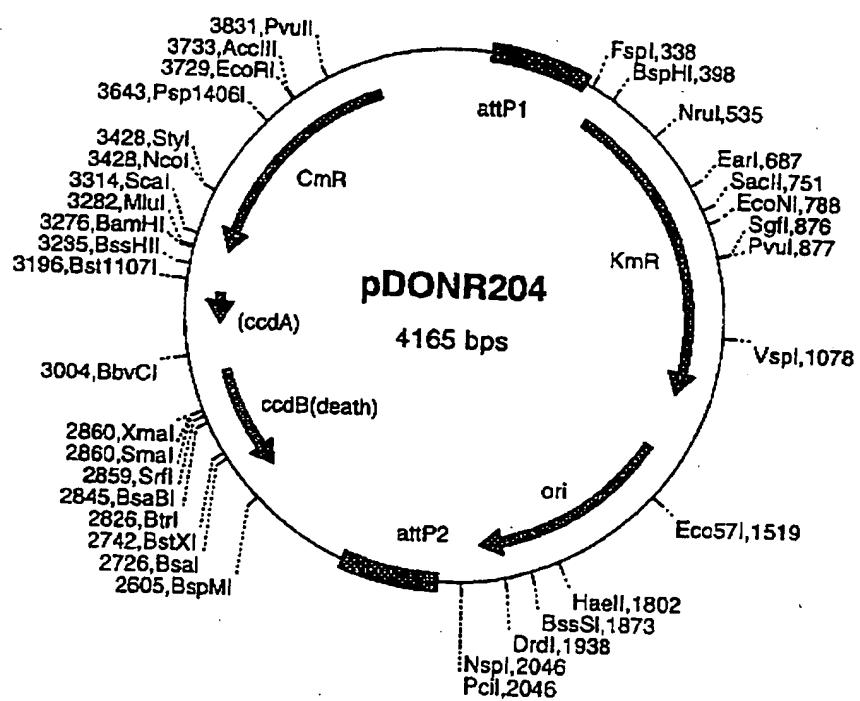
FIGURE 51B

2761 TGGTGAGTAA CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA  
2821 TAAATTCCGT CAGCCAGTTT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC  
2881 CTTGCCATG TTTCAGAAAC AACTCTGGCG CATCGGGCTT CCCATACAAG CGATAGATTG  
2941 TCGCACCTGA TTGCCCAGCA TTATCGCGAG CCCATTTATA CCCATATAAA TCAGCATCCA  
3001 TGTTGGAATT TAATCGCGGC CTCGACGTTT CCCGTTGAAT ATGGCTCATA ACACCCCTTG  
3061 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTCATGA TGATATATT TTATCTTGTG  
3121 CAATGTAACA TCAGAGATT TGAGACACGG GCCAGAGCTG CAGCTAGCAT GGATCTCGGG  
3181 CCCCAAATAA TGATTTTATT TTGACTGATA GTGACCTGTT CGTTGCAACA AATTGATGAG  
3241 CAATGCTTTT TTATAATGCC AACTTTGTAC AAAAAAGCTG AACGAGAAC GTAAAATGAT  
3301 ATAATATATCA ATATATTAAA TTAGATTITG CATAAAAAC AGACTACATA ATACTGTA  
3361 ACACAAACATA TCCAGTCACT ATGAATCAAC TACTTAGATG GTATTAGTGA CCTGTAGTCG  
3421 ACCGACAGCC TTCCAAATGT TCTTCGGGTG ATGCTGCCA CTTAGTCGAC CGACAGCCTT  
3481 CCAAATGTTT TTCTCAAACG GAATCGCTG ATCCAGCCTA CTGCTATTG TCCTCAATGC  
3541 CGTATTAAT CATAAAAAGA AATAAGAAAA AGAGGTGCGA GCCTCTTTT TGTGTGACAA  
3601 AATAAAAACA TCTACCTATT CATATACGCT AGTGTCTAG TCCTGAAAAT CATCTGCATC  
3661 AAGAACAAATT TCACAACCTCT TATACTTTTC TCTTACAAGT CGTTCGGGCTT CATCTGGATT  
3721 TTCAGCCTCT ATACTTACTA AACGTGATAA AGTTTCTGTA ATTTCTACTG TATGACCTG  
3781 CAGACTGGCT GTGTATAAGG GAGCCTGACA TTTATATTCC CCAGAACATC AGGTTAATGG  
3841 CGTTTTGAT GTCATTTCG CGGTGGCTGA GATCAGCCAC TTCTCCCCG ATAACGGAGA  
3901 CGGGCACACT GGCCATATCG GTGGTCATCA TGCGCCAGCT TTCATCCCCG ATATGCACCA  
3961 CGGGTAAAG TTCACGGGAG ACTTTATCTG ACAGCAGACG TGCACTGGCC AGGGGGATCA  
4021 CCATCCGTG CCCGGGCGTG TCAATAATAT CACTCTGTAC ATCCACAAAC AGACGATAAC  
4081 GGCTCTCTCT TTTATAGGTG TAAACCTTAA ACTGCATTTC ACCAGTCCCT GTTCTCGTCA  
4141 GCAAAAGAGC CGTTCATTTCA AATAAACCGG GCGACCTCAG CCATCCCTTC CTGATTTCC  
4201 GCTTTCCA

FIGURE 51C

153/240

FIGURE 52A pDONR204 (kan R)



154/740

## pDONR204 4165 bp

1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTGAG AAGAACATTT  
 61 GGAAGGTGT CGGTGACTA CAGGTCACTA ATACCATCTA AGTAGTTGAA TCATAGTGAC  
 121 TGGATATGTT GTGTTTACA GTATTATGTA GTCTGTTTT TATGCAAAT CTAATTAAAT  
 181 ATATTGATAT TTATATCATT TTACGTTCT CGTTCAGCTT TTTTGTACAA AGTTGGCATT  
 241 ATAAAAAAAGC ATTGCTTATC AATTTGTTGC AACGAACAGG TCACTATCAG TCAAAATAAA  
 301 ATCATTATTT GGGGCCGAG ATCCATGCTA GCTGCAGTGC GCAGGGCCCG TGTCTAAAAA  
 361 TCTCTGATGT TACATTGAC AAGATAAAAA TATATCATCA TGAAACAATAA AACTGTCTGC  
 421 TTACATAAAC AGTAATACAA GGGGTGTTAT GAGCCATATT CAACGGGAAA CGTCTTGCTG  
 481 GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTATAT GGGTATAAAAT GGGCTCGCGA  
 541 TAATGTCGGG CAATCAGGTG CGACAATCTT TCGATTGTAT GGGAAGCCCG ATGCGCCAGA  
 601 GTTGTTCCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT GTTACAGATG AGATGGTCAG  
 661 ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTTTA TCCGTACTCC  
 721 TGATGATGCA TGGTACTCA CCACTGCGAT CCGCAGGAAA ACAGCATTCC AGGTATTAGA  
 781 AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG GCAGTGTTC TGCGCCGGTT  
 841 GCATTGCAATT CCTGTTTGTGTTTAAACAGCGAT CGCGTATTTC GTCTCGCTCA  
 901 GGC GCAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT GATTTTGATG ACGAGCGTAA  
 961 TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATACG CTTTGCCAT TCTCACCGGA  
 1021 TTCAGTCGTC ACTCATGGTG AITTCCTACT TGATAACCTT ATTGTTGACG AGGGAAATT  
 1081 AATAGGTGTGTTGACGAGTCGG AATCGCAGAC CGATACCGAG ATCTTGCAT  
 1141 CCTATGGAAC TGCTCGGTG AGTTTTCTCC TTCAATTACAG AAACGGCTTT TTCAAAATAA  
 1201 TGGTATTGAT AATCCTGATA TGAATAAAATT GCAGTTTCAT TTGATGCTCG ATGAGTTTT  
 1261 CTAATCAGAA TTGGTTAATT GGTTGTAACA CTGGCAGAGC ATTACGCTGA CTTGACGGGA  
 1321 CGGCGNCATG ACCAAAATCC CTTAACGTGA GTTTTCGTT CACTGAGCGT CAGACCCGT  
 1381 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA  
 1441 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTGCGC GATCAAGAGC TACCAACTCT  
 1501 TTTTCCGAAG GTAATGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGT  
 1561 GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT  
 1621 AATCCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC  
 1681 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA  
 1741 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA  
 1801 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG  
 1861 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGTTATCTTT ATAGTCTGT  
 1921 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG  
 1981 CCTATGAAA AACGCCAGCA ACGCGGCCCTT TTACGGTTT CTGGCCTTTT GCTGGCCTT  
 2041 TGCTCACATG TTCTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCTAG  
 2101 CTGGATCGGC AAATAATGAT TTATTTGAT CTGATAGTGA CCTGTTCTGTT GCAACAAATT  
 2161 GATAAGCAAT GCTTTTTAT AATGCCAACT TTGTACAAGA AAGCTGAACG AGAAACGTAA  
 2221 AATGATATAA ATATCAATAT ATTAAATTAG ATTTTGATA AAAAACAGAC TACATAATAC  
 2281 TGTAAAACAC AACATATCCA GTCACTATGA TTCAACTACT TAGATGGTAT TAGTGACCTG  
 2341 TAGTCGACTA AGTTGGCAGC ATCACCCGAC GCACTTGC CGGAATAAAAT ACCTGTGACG  
 2401 GAAGATCACT TCGCAGAATA AATAAACTCT GGTGTCCTG TTGATACCGG GAAGCCCTGG  
 2461 GCCAACTTTT GGGAAAATG AGACGTTGAT CGGCACATTT CACAACCTTT ATACTTTCT  
 2521 CTTACAAGTC GTTCCGGCTTC ATCTGGATTT TCAGCCTCTA TACTTACTAA ACGTGATAAA  
 2581 GTTTCTGTA TTTCTACTGT ATCGACCTGC AGACTGGCTG TGTATAACGG AGCCTGACAT  
 2641 TTATATTCCC CAGAACATCA GGTAAATGGC GTTTTGATG TCATTTCTGC GGTGGCTGAG  
 2701 ATCAGCCACT TCTTCCCCGA TAACGGAGAC CGGCACACTG GCCATATCGG TGGTCATCAT  
 2761 GCGCCAGCTT TCATCCCCGA TATGCACCAAC CGGGTAAAGT TCACGGGAGA CTTTATCTGA  
 2821 CAGCAGACGT GCACTGGCCA GGGGGATCAC CATCCGTCGC CCGGGCGTGT CAATAATATC  
 2881 ACTCTGTACA TCCACAAACA GACGATAACG GCTCTCTCTT TTATAGGTGT AAACCTTAAA  
 2941 CTGCATTCTCA CCACTGCTG TTCTCGTCAG CAAAAGAGCC GTCAATTCA ATAAACCGGG  
 3001 CGACCTCAGC CATCCCTTCC TGATTTCCG CTTTCCAGCG TTCGGCACGC AGACGACGGG  
 3061 CTTCATTCTG CATGGTTGTG CTTACCAGAC CGGAGATATT GACATCATAT ATGCCTTGAG  
 3121 CAACTGATAG CTGTCGCTGT CAACTGTACG TGTAATACGC TGCTTCATAG CACACCTCTT

FIGURE 52B

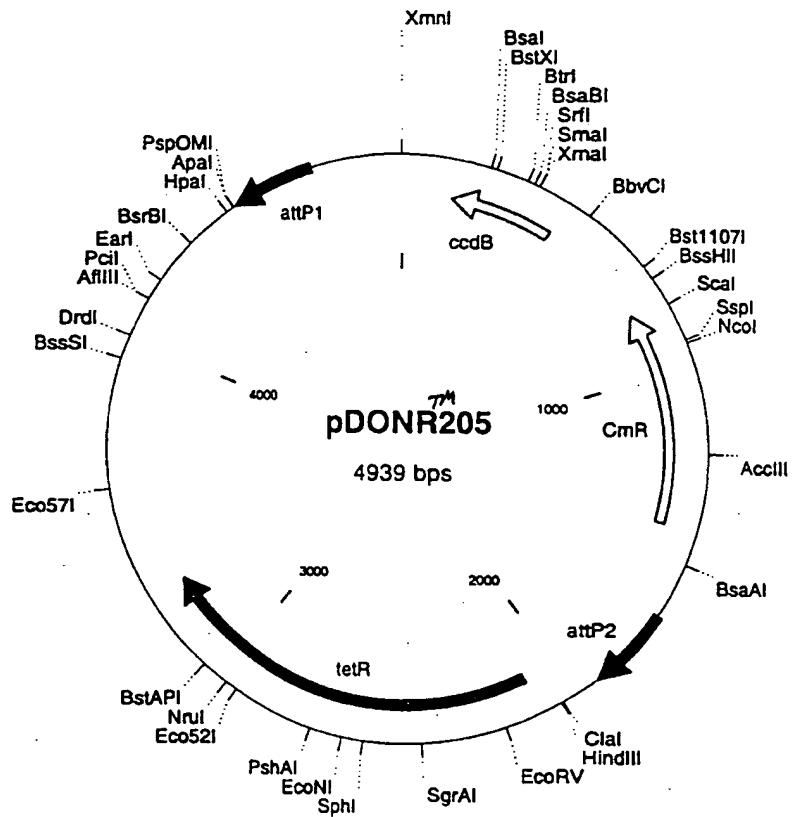
155/240

3181 TTTGACATAC TTGGGTATA CATATCAGTA TATATTCTTA TACCGCAAAA ATCAGCGCG  
3241 AAATACGCAT ACTGTTATCT GGCTTTAGT AAGCCGGATC CACCGGTTA CGCCCCGCC  
3301 TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTG TGCCGACATG GAAGCCATCA  
3361 CAGACGGCAT GATGAACCTG AATCGCCAGC GGCATCAGCA CCTTGTGCC TTGCGTATAA  
3421 TATTTGCCCA TGGTGAAAAC GGGGGCGAAG AAGTTGTCCA TATTGGCCAC GTTTAAATCA  
3481 AAACTGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCT  
3541 TTAGGGAAAT AGGCCAGGTT TTCACCGTAA CACGCCACAT CTTGCGAATA TATGTGTAGA  
3601 AACTGCCGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA  
3661 TGGAAAACGG TGTAACAAGG GTGAACACTA TCCCATATCA CCAGCTCACC GTCTTTCATT  
3721 GCCATACCGA ATTCCGGATG AGCATTCACTC AGGCGGGCAA GAATGTGAAT AAAGGCCGA  
3781 TAAAACTTGT GCTTATTTTTT CTTTACGGTC TTTAAAAGG CCGTAATATTC CAGCTGAACG  
3841 GTCTGGTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAAATGTTT TTTACGATGC  
3901 CATTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCCTTA  
3961 GCTCCTGAAA ATCTCGATAA CTCAAAAAT ACGCCCGGT A GTGATCTTAT TTCATTATGG  
4021 TGAAAGTTGG AACCTCTTAC TGTTCTTGAT GCAGATGATT TTCAGGACTA TGACACTAGC  
4081 ATATATGAAT AGGTAGATGT TTTTATTTG TCACACAAAA AAGAGGCTCG CACCTTTTT  
4141 TCTTATTTCT TTTTATGATT TAATA

FIGURE 52C

156/240

Figure S3A: pDONR205 (tetR)



157/240

## pDONR205 4939 bp

GGCATCAGCACCTTGTGCCCTGCGTATAATATTGCCATGGTAAAAACGGGGCGAAG  
 AAGTTGTCCATATTGCCACGTTAAATCAAACGGTAAACTCACCCAGGGATTGGCT  
 GAGACGAAAACATATTCTCAATAAACCCCTTAGGGAAATAGGCCAGGTTTACCGTAA  
 CACGCCACATCTTGCAGATATATGTGTAGAAACTGCCGAAATCGTCGTTATTCACTC  
 CAGAGCGATGAAAACGTTTCACTGGAAAACGGTGTACAAGGGTGAACACTA  
 TCCCATACTACCAGCTACCGTCTTCATTGCCATACGGAATTCCGGATGAGCATTCA  
 AGGGGGCAAGAATGTGAATAAAGGCCGATAAAACTTGTGCTTATTTCCTTACGGTC  
 TTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGAC  
 TGAAATGCCCTAAAATGTTCTTACGATGCCATTGGGATATATCAACGGTGGTATATCCA  
 GTGATTTTTCTCCATTAGCTCCTAGCTCCTGAAAATCTCGATAACTCAAAAAT  
 ACGCCCGGTAGTGTATTTCTTACAGTATTATGTAGTCTGTTTATGCAAATCTAA  
 ACGTCTCATTTGCCAAAAGTTGCCAGGGCTCCGGTATCAACAGGGACACCCAGGA  
 TTATTTATCTCGCAAGTGATCTCCGTACAGGTATTIATTGGCGCAAAGTGCCTCG  
 GGTGATGCTGCCAACCTAGTCGACTACAGGTACTAATACCATCTAAGTAGTTGATTCA  
 AGTGACTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTTTATGCAAATCTAA  
 TTTAATATATTGATATTATATCATTTACGTTCTCGTCAGCTTCTGTACAAAGTT  
 GGCATTATAAGAAAGCATTGCTTATCAATTGTTGCAACGAACAGGTCACTATCAGTC  
 AATAAAAATCATTATTGCCATCCAGCTGCAGCTCTGGCCCGTGTCTAAAATCTGATG  
 TTACATTGACAAGATAAAAATATCATCATGAATTCTCATGTTGACAGCTTATCATC  
 GATAAGCTTAATGCCGTTAGTTATCACAGTTAAATTGCTAACGCAGTCAGGACCGTGT  
 ATGAAATCTAACATGCCCTCATCGTCATCCTGGCACCGTCACCCCTGGATGCTGTAGGC  
 ATAGGCTGGTTATGCCGGTACTGCCGGCTTGTGGGATATGTCATTCCGACAGC  
 ATCGCCAGTCACTATGCCGCTGCCAGCGTATATGCGTTGATGCAATTCTATGCGCA  
 CCCGTTCTCGGAGCAGTCCGACCGCTTGGCCGCCAGCTCTGCTCGCTCGCTA  
 CTTGGAGCCACTATCGACTACCGGATCATGGCACCCGTCACCCCTGGATGCTGTAGGC  
 GCCGGACGCATCGTGGCCGCACTACCGGGGCCACAGGTGCGGTTGCTGGCGCTATATC  
 GCCGACATACCGATGGGAAGATGGGCTGCCACTTCGGGCTCATGAGCGCTTGT  
 GGCGTGGGTATGGTGGCAGGCCGGTGGCGGGGACTGTTGGCGCCATCTCTGCT  
 GCACCATTCCTTGCAGGGGGGGTGTCAACGGCTCAACCTACTACTGGGCTGCTTCTA  
 ATGCAAGGAGTCGATAAGGGAGAGCGTCAGCGATGCCCTGAGAGCCTTCAACCCAGTC  
 AGCTCCTTCCGGTGGGCGGGGCGATGACTATCGTCGCCACTTATGACTGTCTTCTT  
 ATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAATTTCGGCGAGGACCGC  
 TTGCGCTGGAGCGCAGGATGATCGGCTGTGCTTGTGGTATTGGAAATCTGCAAGCC  
 CTCGCTCAAGCTTCTGCACTGGTCCGCCACAAACGTTGGCGAGAAGCAGGCCATT  
 ATCGCCGGCATGGCGGCCAGCGCTGGGCTACGTTGCTGGGTTGCGACCGAGGC  
 TGGATGGCTTCCCCATTATGATTCTTCGCTTCCGGGGCATGGGATGCCCGCTTG  
 CAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTCAAGGATCGCTC  
 GCGGCTTCTACCAGCTAACGGTCAACTTCGATCATGGACCGCTGATCGTCACGGGATTTATGCC  
 GCCTCGGGAGCACATGGAACGGGTTGGCATGGATTGAGGGCCGCTATACCTGTC  
 TGCCCTCCCCCGGTTGCGTGGGTGCACTGGAGCCGGGCCACCTGACCTGAATGGAAAGCC  
 GGCGGACCTCGCTAACGGATTACCAACTCCAGAATTGGAGCAATCAATTCTGCGGA  
 GAACGTGAATGCGAACCAACCTTGGCAGAACATATCCATCGCATGACCAAAATCCC  
 TTAACGTGAGTTTCTGCTTCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGATCTTC  
 TTGAGATCTTTCTGCGCTAATCTGCTGCTTGTCAAACAAAAAAACCCAGCTTAC  
 AGCGGTGGTTGTTGCCGGATCAAGAGCTACCAACTCTTCTCGAAGGTAACGGCTT  
 CAGCAGAGCGCAGATAACCAAATACTGTCCTCTAGTGTAGCGTAGTTAGGCCACCACTT  
 CAAGAACTCTGAGCAGCCCTACATACCTCGCTCTGCTAATCTGTTACCGAGTGGCTGC  
 TGCCAGTGGCAGATAAGTCGTCTTACGGGTTGGACTCAAGACGATAGTTACCGGATAA  
 GGCGCAGGGTGGCTGAACGGGGGTTGCGCACACAGCCAGCTGGAGCGAACGAC  
 CTACACCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCACCGCTTCCGAAGG  
 GAGAAAGGGGGACAGGTATCCGGTAAGCGGCAGGGTGGAAACAGGAGAGCGCACGAGGGA  
 GCTTCCAGGGGAAACGCCCTGGTATCTTATAGTCCTGTCGGGTTTGGCCACCTGACT  
 TGAGCGTCACTTGTGATGCTGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAA-

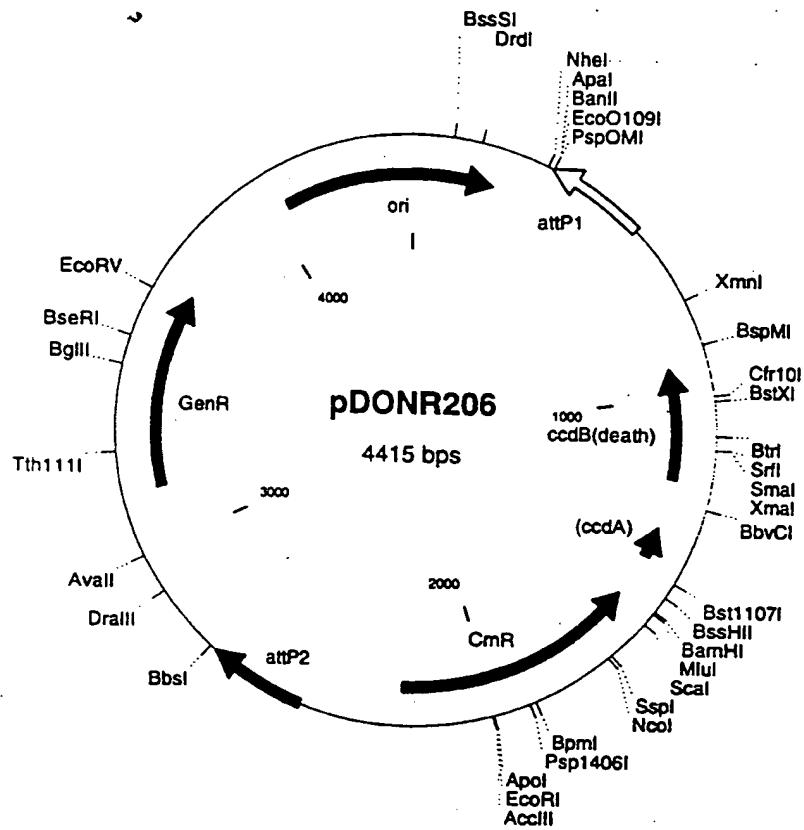
FIGURE 53B

158/240

CGCGGGCTTTTACGGTTCTGGCCTTTGCTGGCTTTGCTCACATGTTCTTCCTGC  
GTTATCCCTGATTCTGTGGATAACCGTATTACCGTAGGCCAGGAAGAGTTGTAGAAC  
GCAAAAAGGCCATCCGTAGGATGGCCTTCTGCTTAGTTGATGCCCTGGCAGTTATGGC  
GGCGTCTGCCCCGCCACCCCTCCGGGCGTTGCTTACAACGTTCAAATCCGCTCCCGC  
GGATTTGTCTACTCAGGAGAGCGTTCACCGACAAACAAACAGATAAAACGAAAGGCCAG  
TCTTCCGACTGAGGCCCTTGTGTTTATTTGATGCCCTGGCAGTTCCCTACTCTCGCTTAAC  
GCTAGCATGGATCTCGGGCCCAAATAATGATTTATTTGACTGATAGTGACCTGTCG  
TTGCAACAAATTGATGAGCAATGCTTTTATAATGCCAACTTTGACAAAAAGCTGAA  
CGAGAACAGTAAAAATGATATAAATATCAATATATTAAATTAGATTTGCATAAAAACAG  
ACTACATAAATCTGAAACACACATATCCAGTCACATGAACTCAACTACTAGATGGT  
ATTAGTGACCTGTAGTCGACCGACAGCCTTCAAATGTTCTCGGGTGATGCTGCCACT  
TAGTCGACCGACGCCCTCAAATGTTCTCTAAACGGAATCGTCGTATCCAGCCTACT  
CGCTATTGTCCTCAATGCCGTATTAATCATAAAAAGAAATAAGAAAAGAGGTGCGAGC  
CTCTTTTTGTGTGACAAAATAAAACATCTACCTATTCAATACGCTAGTGTCAAGTC  
CTGAAAATCATCTGCATCAAGAACAACTTCAAACTTTACTTTCTTTACAAGTCG  
TTCGGCTTCATCTGGATTTTCAAGCCTCTATACTTACTAAACGTGATAAAGTTCTGTAAT  
TTCTACTGTATCGACCTGCAGACTGGCTGTGATAAGGGAGCCTGACATTATATTCCCC  
AGAACATCAGGTTAATGGGTTTTGATGTCAATTTCGCGGTGGCTGAGATCAGCCACTT  
CTTCCCCGATAACGGAGACCGGCACACTGGCATATCGGTGGTCATCATGCGCAGCTT  
CATCCCCGATATGCACCACCGGGTAAAGTTCAAGGGAGACTTTATCTGACAGCAGACGTG  
CACTGGCAGGGGATCACCATCGTCGCCGGCGTGTCAATAATATCACTCTGTACAT  
CCACAAAACAGCAGATAACGGCTCTCTTTATAGGTGTAACCTTAAACTGCATITCAC  
CAGTCCCTGTTCTCGTCAGCAAAAGAGCCGTTCAATTCAATAACCGGGGACCTCAGCC  
ATCCCCTCCTGATTTCCGCTTTCCAGCGTTGGCACGCAGACGACGGGCTTCATTCTGC  
ATGGTTGTGCTTACCAAGACGGAGATATTGACATCATATGCTTGAGCAACTGATAGC  
TGTCGCTGTCAACTGTCACTGTAATACGCTGTTCAAGCACACCTTTGACATACT  
TCGGGTATAACATATCAGTATATAATTCTTATACCGAAAAATCAGCGCAGAACACGATA  
CTGTTATCTGGCTTTAGTAAGCCGGATCCACCGCATTACGCCCGCCCTGCCACTCATC  
GCAGTACTGTGTAATTCAATTACGCACTGCGACATGGAAGCCATCACAGACGGCATG  
ATGAACCTGAATGCCAGC

FIGURE 53C

159/240



160/260

## pDONR206 4415 bp

CGGCATTGAGGACAATAGCGAGTAGGCTGGATACGACGATTCCGTTGAGAAGAACATT  
 GGAAGGCTGTCGGTCGACTACAGGTCACTAATACCCTCTAAGTAGTTGAATCATAGTGAC  
 TGGATATGTTGTTTACAGTATTATGTTAGCTGTTTATGCAAATCTAATTAAAT  
 ATATTGATATTATCATTACAGTCTCGTCAGCTTTGTACAAAGTTGGCATT  
 ATAAAAAAAGCATTGCTTATCAATTGTTGCAACGAACAGGTCACTATCAGTCAAATAAA  
 ATCATTATTGGGCCGAGATCCATGCTAGCGGTAAACCGTTATCCACAGAACAGGG  
 GATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAG  
 GCCCGTTGCTGGCTTTCCATAGGCTCCGCCCTGACGAGCATTACAAAAATCGA  
 CGCTCAAGTCAGAGGTTGGCGAAACCCGACAGGACTATAAGATACCAAGGCCTTCCCCCT  
 GGAAGCTCCCTCGTGCCTCTCGTCCGACCCCTGCCGTTACCGGATACCTGTCCGCC  
 TTCTCCCTCGGAAAGCGTGGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTCG  
 GTGTAGGTCGTTGCTCAAGCTGGCTGTGTCAGAACCCCCCGTTGACCCGACCCG  
 TGCGCTTATCCGTAACATCGTCTGAGTCCAACCCGTAAGACACGACTTATCGCCA  
 CTGGCAGCAGCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG  
 TTCTGAAAGTGGTGGCTAACTACGGCTACACTAGAACGGAGTATTTGGTATCTCCGCT  
 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTGGTAGCTCTGATCCGGAAACAAACC  
 ACCGCTGGTAGCGGTGGTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA  
 TCTCAAGAAGATCCTTGATCTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCA  
 CGTTAAGGGATTTGGTCAATGNCGGCTCCGCTAACGTCAAGCTGTAACTGCCAGTGT  
 TACAACCAATTAAACCAATTCTGATTAGAAAACCTATCGAGCATCAAATGAAACTGCAAT  
 TTATTCAATACAGGATTATCAATACCATATTGGAAAAGCCGTTCTGTAATGAAGGA  
 GAAAACCTACCGAGGCAGTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCG  
 ACTCGTCAACATCAATACAACCTATTAGCCAGGTCTCCGATCTCTGAGGCCAGGGC  
 AGATCCGTGACAGCACCTTGCCGTTAGAACAGCAAGGCCAGCAATGCCCTGAGGATGC  
 GTGGAGACCGAAACCTTGCGCTCGCCAGCCAGGACAGAAATGCCCTGACTTCGCTG  
 CTGCCAAGGTTGCCGGGTGACGACACCGTGGAAACGGATGAAGGCAAGAACCCAGTTG  
 ACATAAGCCTGTTGGTCTGAAACTGTAATGCAAGTAGCGTATGCCCTCACGCAACTGG  
 TCCAGAACCTTGACCGAACGCACTTGCCGTTAGAACGGCGAGTGGCGTTTCTGGCTTGT  
 TATGACTTTTTGTTGATGTTATGGAGCAGCAACGATGTTACGCAAGCACGATGTTAC  
 GCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTAGTGGCATCATTGCA  
 ATGTTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGCTCTTGTATCTGGCTTGT  
 TGAGTTGGAGACGTTAGCCACCTACTCCAAACATGAGGGACTCCGATTACCTGGGAA  
 CTTGCTCGTAGTAAGACATTACGCGCTTGTGCTGCCCTGACGAGAACAGGGTTGG  
 CGCTCTCGGGCTTACGTTCTGCCAGGTTGAGCAGCCGCTAGTGAGATCTATATCTA  
 TGATCTCGAGTCCTCGGCGAGCACGGAGGAGGGCATTGCCACCGCGCTCATCAATCT  
 CCTCAAGCATGAGGCCAACCGCCTGGTCTTATGTGATCTACGTCAGCAAGCAGATTACGG  
 TGACGATCCGCAGTGGCTCTCTATACAAAGTTGGCATACGGGAAGAAGTGTGACTT  
 TGATATGACCCAAGTACCGCCACCTAACAACTCGTCAAGCCGAGATCGGCTTCCGGC  
 CTAATTTCCCTCGTAAAAAAATAGGTTATCAAGTGAGAAATCACCATGAGTGACCTG  
 AATCCGGTGAGAATGGCAAAAGCGTATGCAATTCTTCCAGACTTGTCAACAGGCCAGC  
 CATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCAATTGTAATTGCG  
 CCTGAGCGAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAACAGGAATCGAAT  
 GCAACCGGGCAGGAAACACTGCCAGCGCATCAACAAATTTCACCTGAATCAGGATATT  
 CTTCTAATACCTGGATGCTGTTCCCGCGATCGCAGTGGTAGTAACCATGATCAT  
 CAGGAGTACGGATAAAATGCTGATGGTCGGAAGAGGCAATAATTCCGTCAGCCAGTTA  
 GTCTGACCATCTCATGTAACATCATTGGCAACGCTACCTTGCCATGTTCAGAAACA  
 ACTCTGGCGCATCGGCTTCCCATACAATCGAAAGATTGTCGCACCTGATTGCCGACAT  
 TATCGCGAGCCCATTATACCCATATAATCAGCATCCATGTTGAAATTAAATCGCGGCC  
 TCCAGCAAGACGTTCCGTTGAATGGCTCATAACACCCCTGATTACTGTTATGT  
 AAGCAGACAGTTTATTGTTCATGATGATATAATTAAATCTTGTGCAATGTAACATCAGA  
 GATTTGAGACACGGGCCNGCGCACTGCGAGCTGGATGGCAAAATAATGATTTTATTG  
 ACTGATAGTGCACCTGTTGCAACAAATTGATAAGCAATGCTTTTATAATGCCAAC -

FIGURE 54B

(61 / 240)

TTTGTACAAGAAAGCTGAACGAGAAACGTAAAATGATATAAATATCAATATATTAAATTA  
GATTTGCATAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCACATAG  
ATTCAACTACTTAGATGGTATTAGTGACCTGTAGTCGACTAAGTTGGCAGCATACCGA  
CGCACTTGCAGCGAATAAAATACCTGTGACGGAAGATCACTTCGCAGAATAAAATCC  
TGGTGTCCCTGTTGATACCGGGAGGCCCTGGCCAACCTTTGGCGAAATGAGACGTTGA  
TCGGCACGTAAGAGGTTCAACTTACCCATAATGAAATAAGATCACTACCGGGCGTATT  
TTTGAGTTATCGAGATTTCAAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGG  
ATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTGAGGCATTCAGTC  
AGTTGCTCAATGTACCTATAACCAGACCGTTGAGCTGGATATAACGGCTTTAAAGAC  
CGTAAAGAAAAATAAGCACAAGTTTATCCGGCTTTATCACATCTTGCCCGCTGAT  
GAATGCTCATCCGGATTCCGTATGGCAATGAAAGACGGTAGCTGGTATATGGGATAG  
TGTTCACCCCTGTTACACCGTTTCCATGAGCAAACGTAAACGTTTTCATCGCTCTGGAG  
TGAATACCAACGACGATTCCGGCAGTTCTACACATATATTCGCAAGATGTGGCGTGT  
CGGTGAAACCTGGCTTATTCCTAAAGGTTATTGAGAATATGTTTTCGTCTCAGC  
CAATCCCTGGGTGAGTTTACCAAGTTTGTGATTTAACTGGCCAATATGGACAACCTCTT  
CGCCCCGGTTTTCACCATGGGAAATAATTAACGCAAGGGCACAAGGTGCTGATGCCGCT  
GGGATTCAGGTTCATCATGCCGCTGTGATGGCTTCCATGCGGAGAATGCTTAATGA  
ATTACACAGTACTGCGATGAGTGGCAGGGGGGGCGTAAACGCGTGGATCCGGCTTACT  
AAAAGCCAGATAACAGTATGCGTATTTGCGCCTGATTTTGCCTATAAGATAATAC  
TGATATGTATAACCGAAGTATGTCAAAAGAGGTGTGCTATGAGCAGCTTAACTG  
ACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATGATGTCAATATCTCCGGTC  
TGGTAAGCACAACCATGCGAGAATGAAAGCCGCTGCTGCGTGGCAACGCTGGAAAGCGG  
AAAATCAGGAAGGGATGGCTGAGGTCGCCGGTTTATTGAAATGAAACGGCTTTTGTG  
ACGAGAACAGGGACTGGTGAATGCGATTAAAGGTTTACACCTATAAAAGAGAGAGCCGT  
TATCGTCTGTTGTGGATGTACAGAGTGTGATATTATTGACACGCCGGCAGGGATGGTG  
ATCCCCCTGGCCAGTGCACGTCGCTGTCAGATAAGTCTCCCGTGAACCTTACCCGGTG  
GTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTC  
TCCGTTATCGGGGAGAACGACTGGCTGATCTCAGCCACCGGAAATGACATCAAAACGCC  
ATTAACCTGATGTTCTGGGAATATAATGTCAGGCTCCGTTATACACGCCAGTCTGCA  
GGTCGATACAGTAGAAATTACAGAAACCTTATCACGTTAGTAAGTATAGAGGCTGAAAA  
TCCAGATGAAGCCGAAACGACTTGTAAAGAGAAAAGTATAAGAGTTGTGAAATTGTTCTG  
TGCAGATGATTTCAAGGACTATGACACTAGCATATATGAATAGGTAGATGTTTTTATTT  
GTCACACAAAAAGAGGCTCGCACCTTTTCTTATTTCTTATGATTTAATA

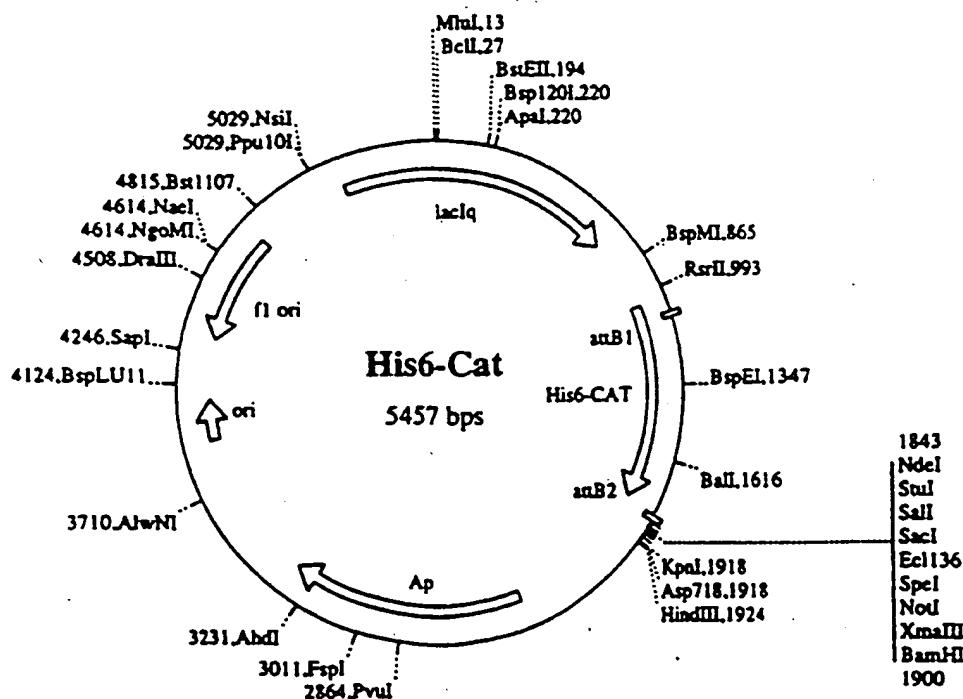
FIGURE 54 C

162 / 240

**Figure 55** An Entry (pEMR7) Clone of CAT Subcloned into pDEST2

TEV protease → Start CAT

1123	Tyr	Phe	Gln ↓ Gly	Thr	Met	Gly	Lys	Lys	Ile	Thr	Gly	Tyr	Thr	Thr	Val	Arg	
	tat	ttt	caa	ggd	acc	atg	gag	aaa	aaa	atc	act	gga	tat	acc	acc	gtt	gat
	ata	aaa	gtt	cct	tgg	tac	ttc	ttt	tag	tga	cct	ata	tgg	tgg	caa	cta	



163/240

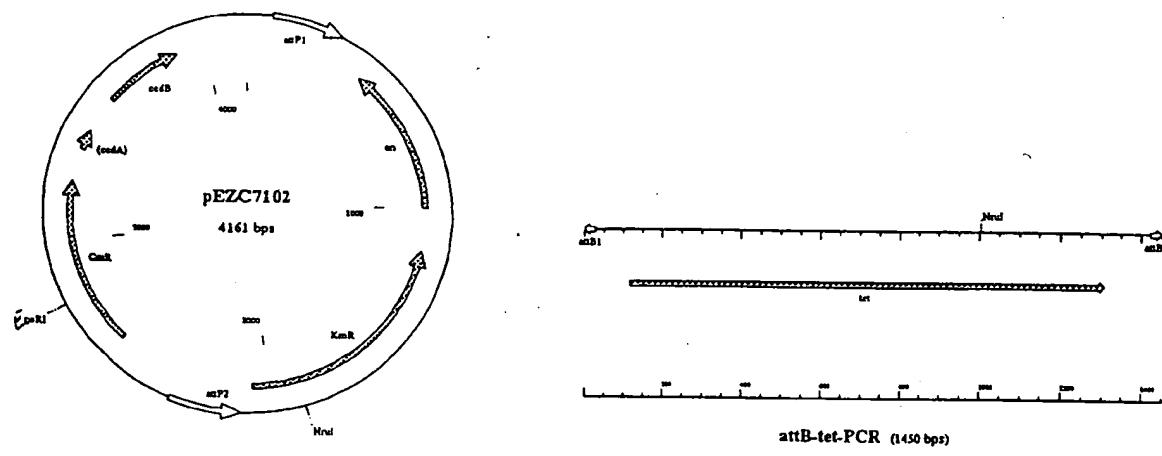


FIGURE 56

164/240

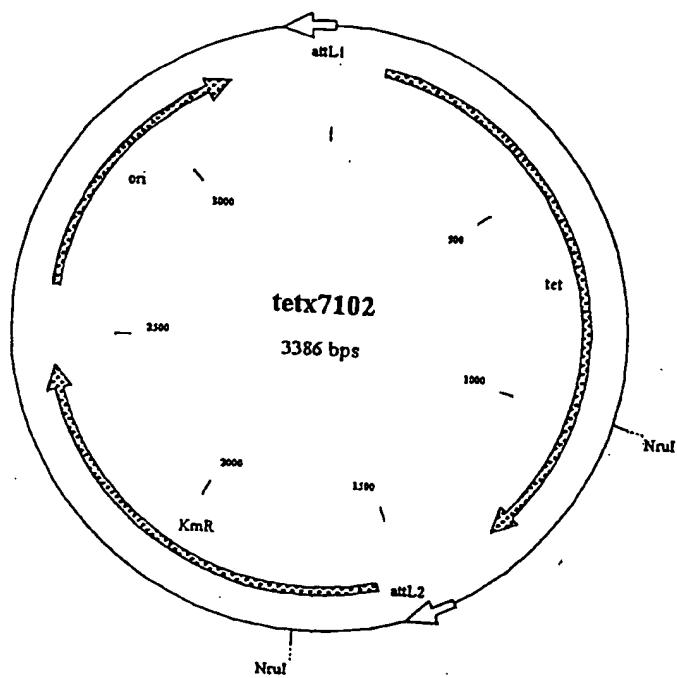


FIGURE 57

165/240

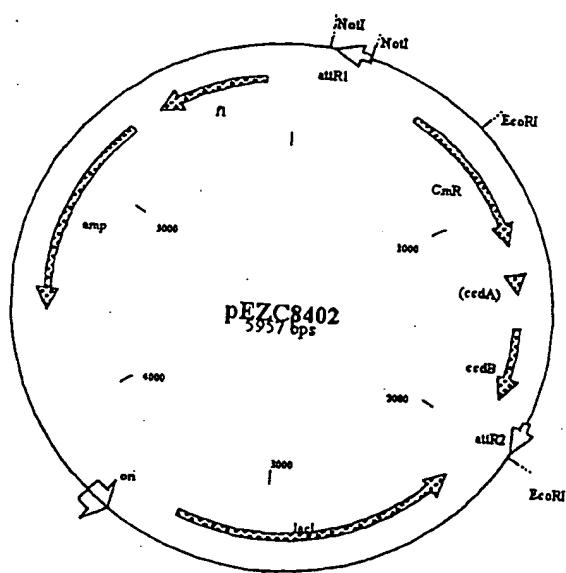


FIGURE 58

166/260

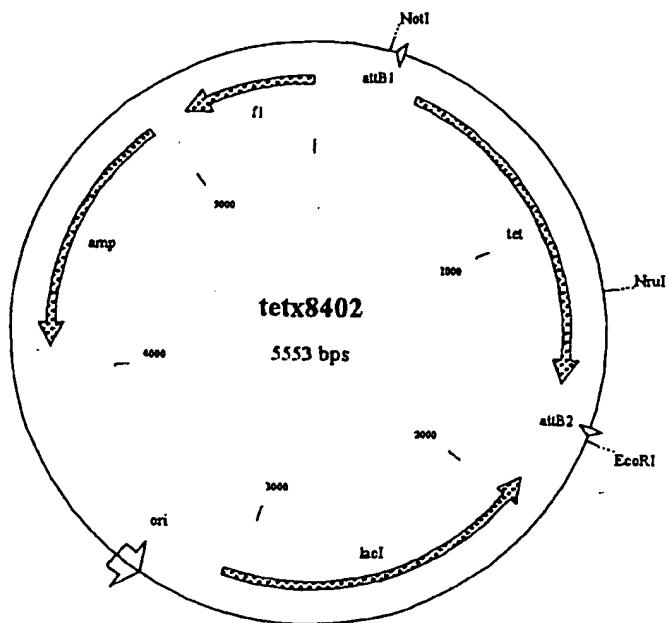


FIGURE 59

167/240

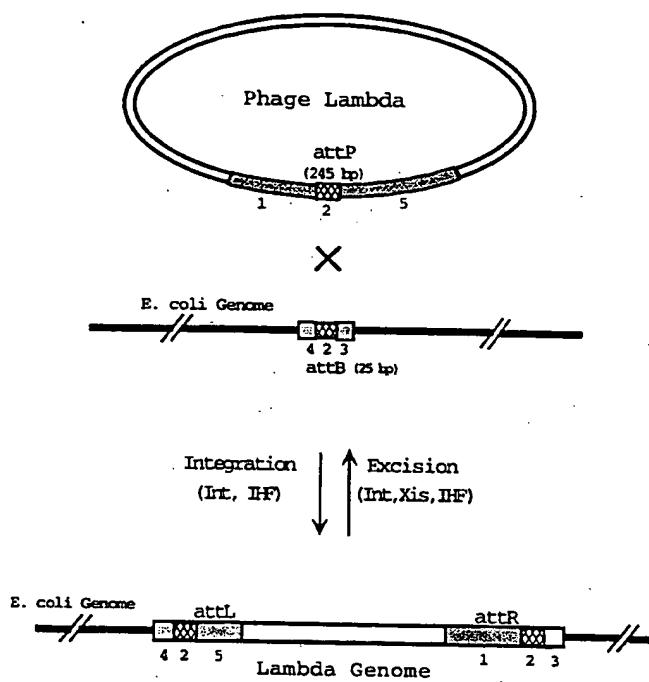


FIGURE 60

168/240

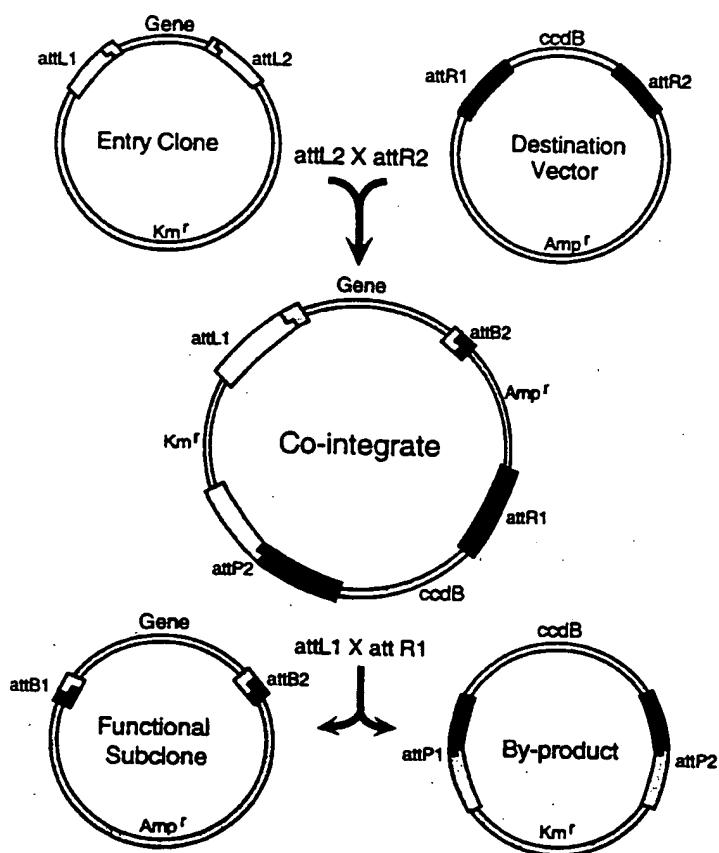


FIGURE 61

169/240

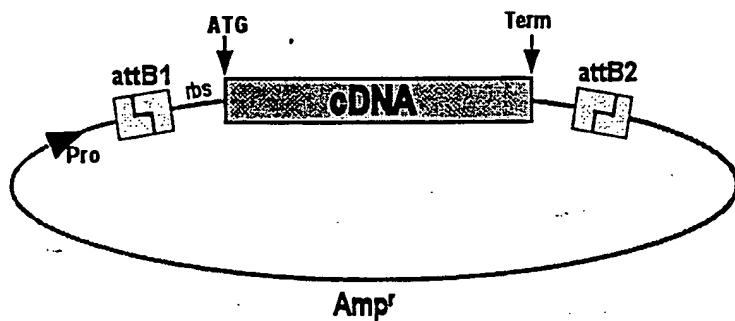
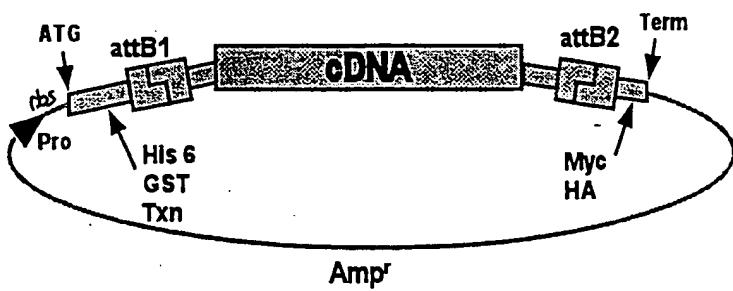
**Native Protein Expression:****Fusion Protein Expression:**

FIGURE 62

170/240

Mlu I (reading frame A)

Bgl II (reading frame B)

Xba I (reading frame C)

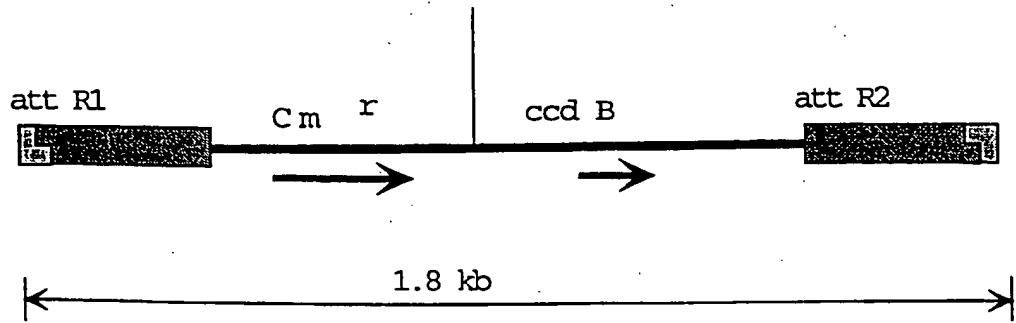
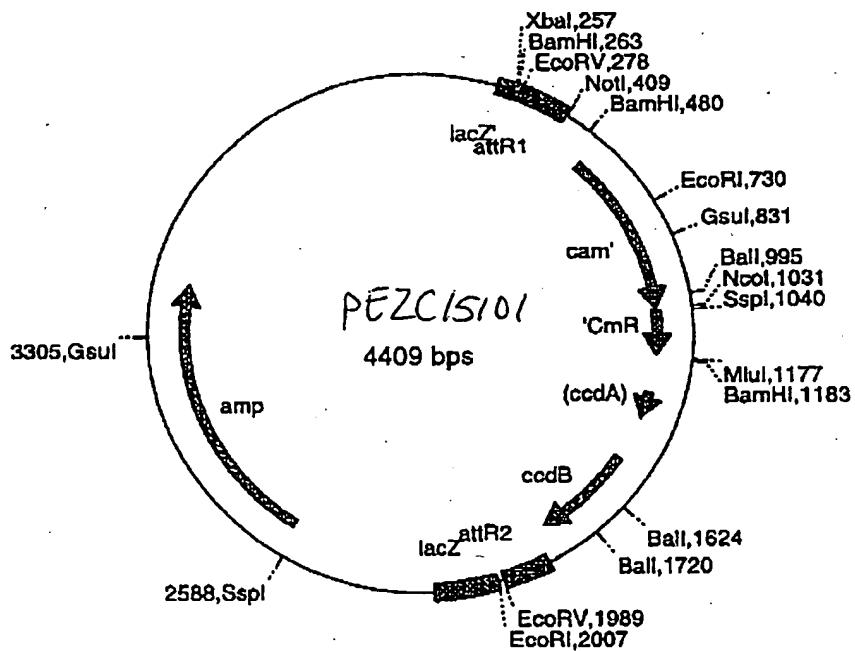


FIGURE 63

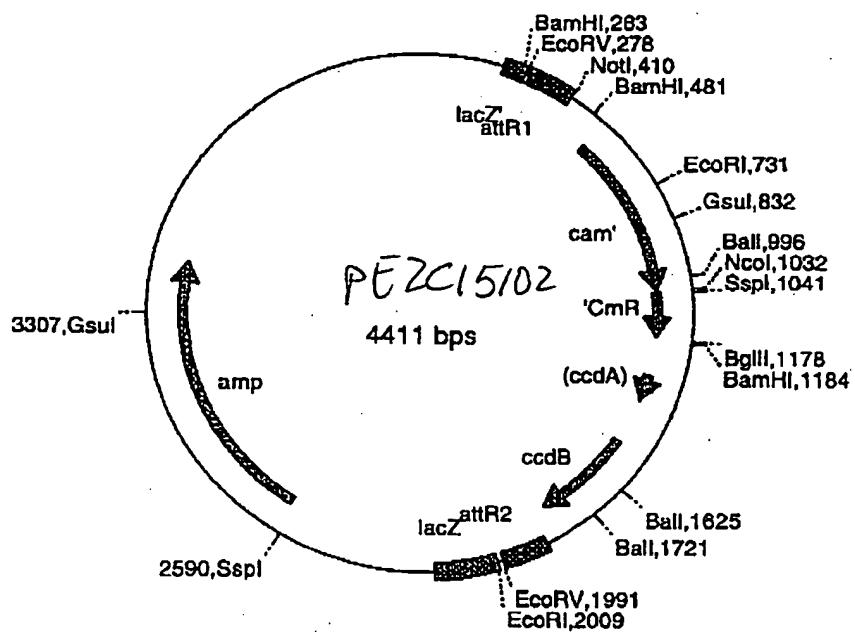
171/240

FIGURE 64A



172/240

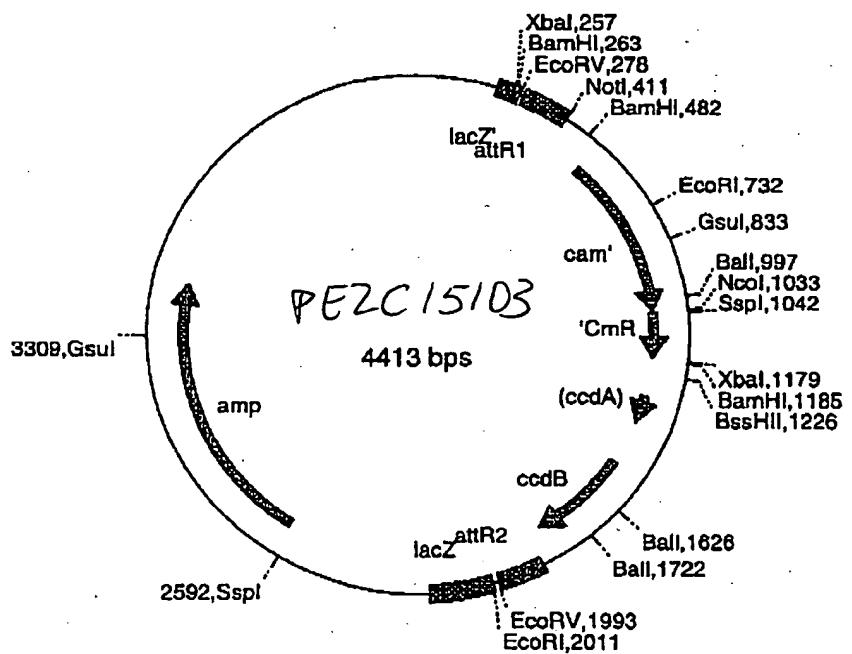
FIGURE 6aB



173/260

5

FIGURE 64C



174/260

Primers for Amplifying *tet<sup>R</sup>* and *amp<sup>r</sup>*  
for Cloning by Recombination

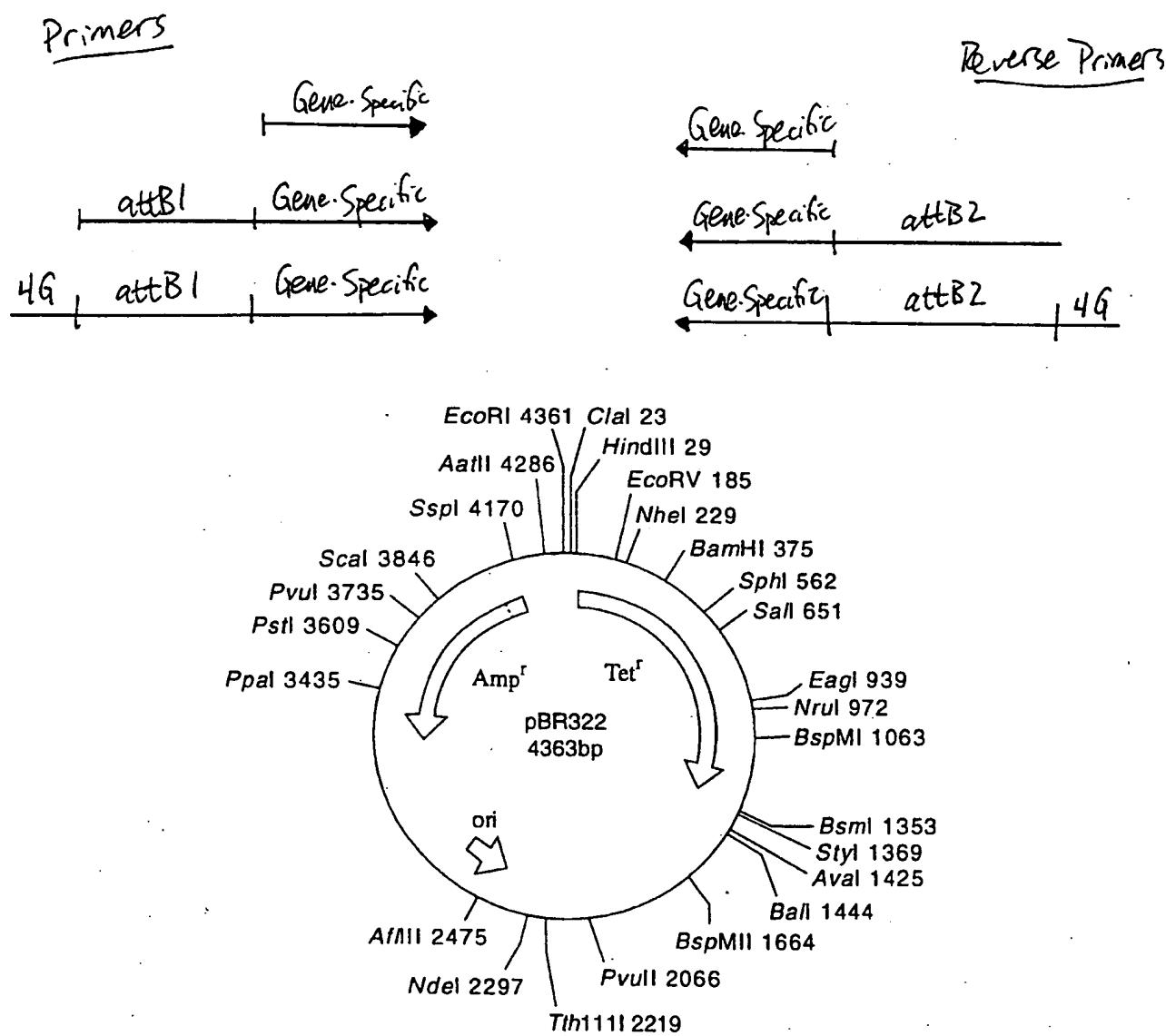


FIGURE 65

175/260

**Results of Cloning  
tet and amp PCR Products  
by Recombination**

<b>PCR Product Used in GCS Reactions</b>	<b>No. Colonies Obtained (100 µl plated)</b>	<b>Form of DNA Analyzed</b>	<b>Colonies Obtained of Predicted Size</b>
tet	6, 10	SC	0 of 8
attB-tet	9, 6	SC	1 of 8
attB+4G-tet	824, 1064	SC AvaI+Bam	7 of 7 7 of 7
amp	7, 13	SC	0 of 8
attB-amp	18, 22	SC	3 of 8
attB+4G-amp	3020, 3540	SC PstI	8 of 8 8 of 8
attB Plasmid (Pos. Control)	320, 394		

FIGURE 66

176/260

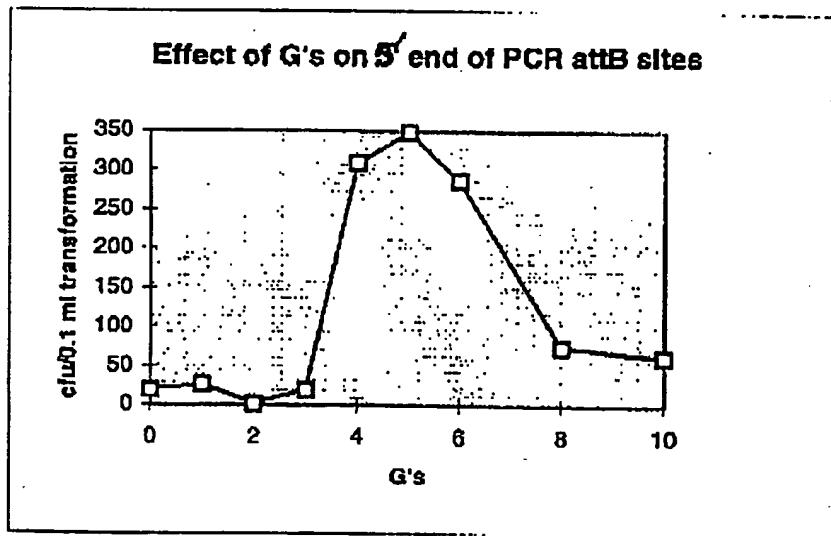


FIGURE 67

177/240

**Titration of attP amounts with various attB PCR amounts in  
 $B \times P$  Reaction**

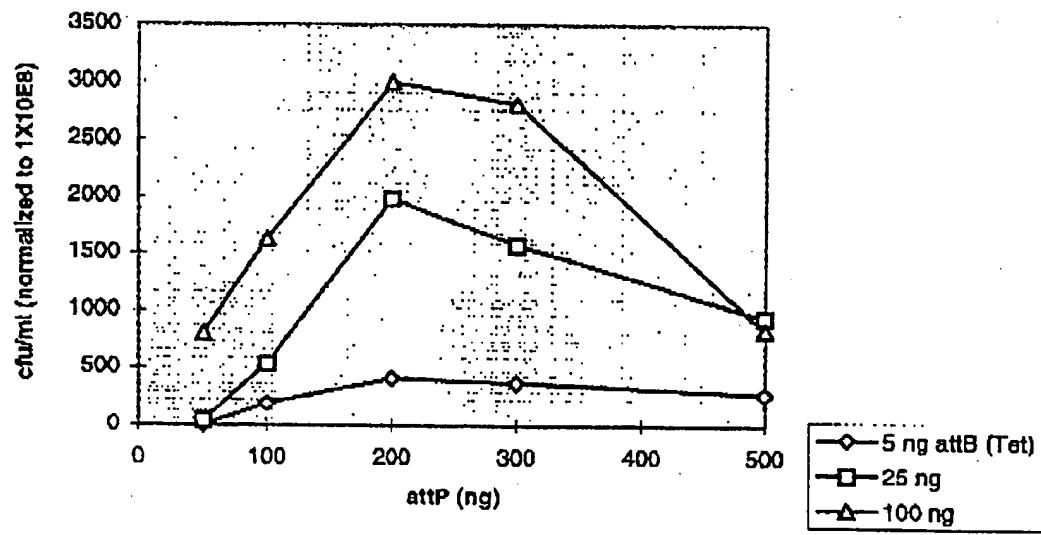
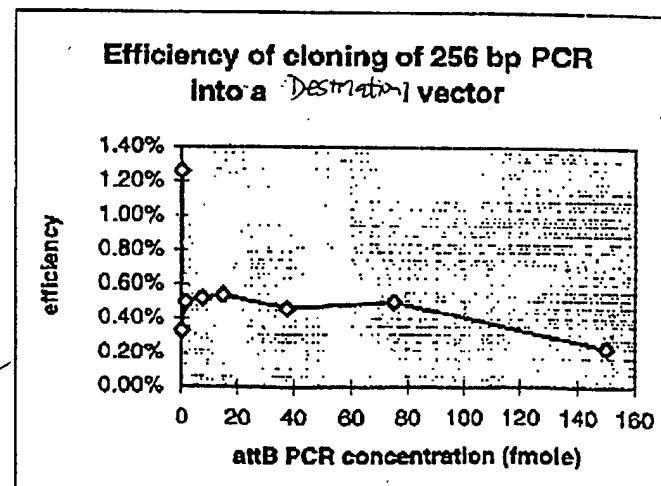
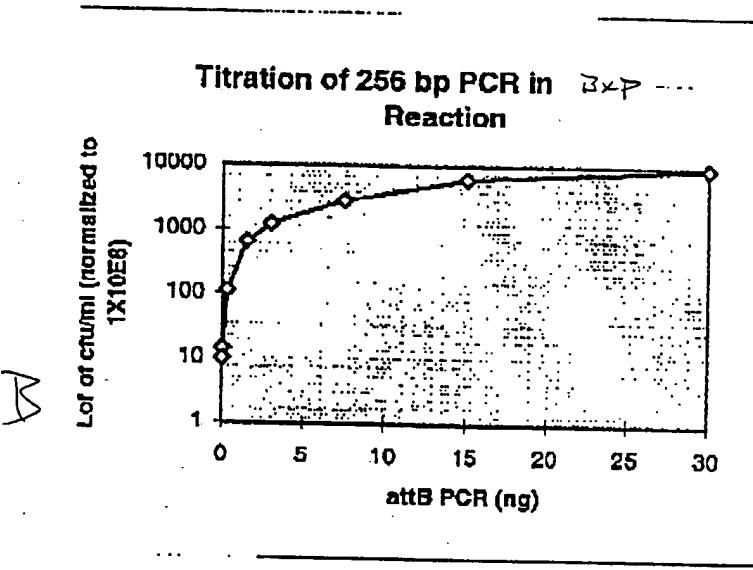
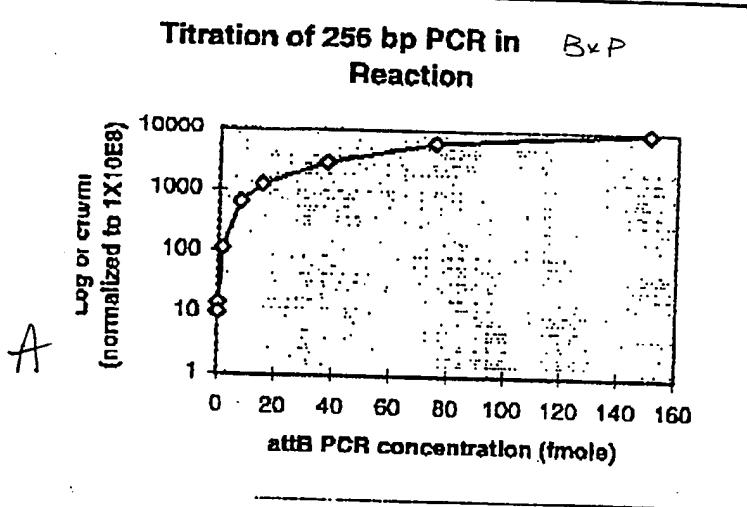


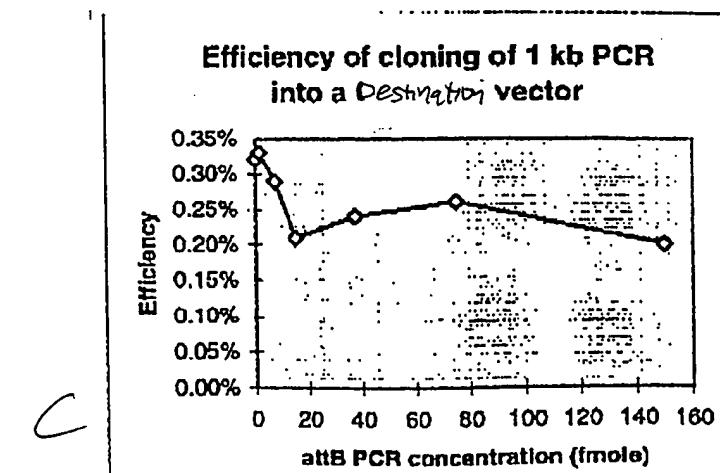
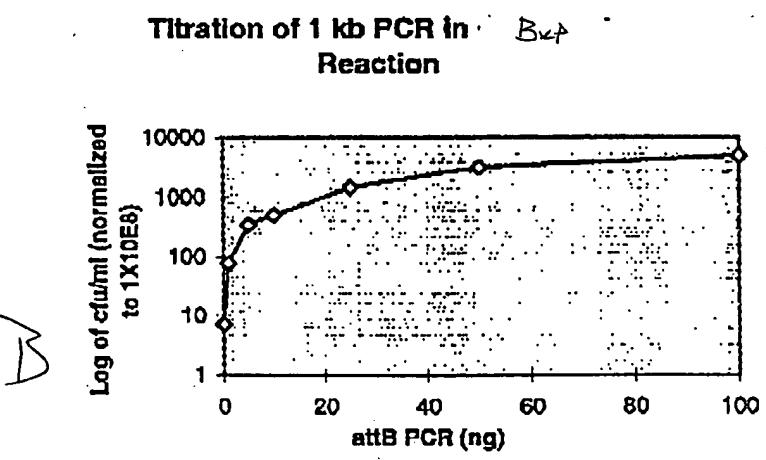
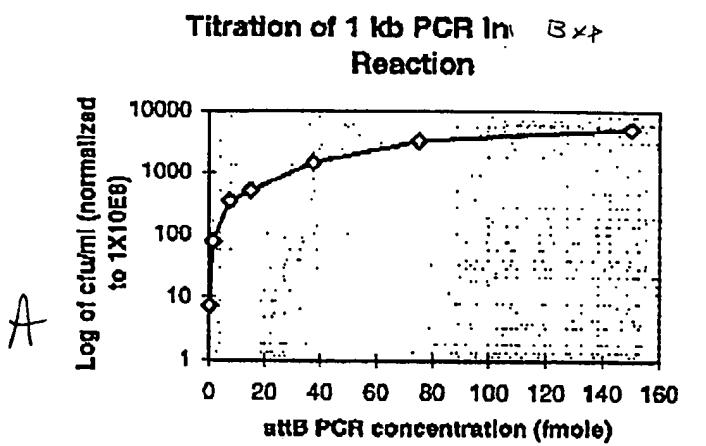
FIGURE 68

178/250

FIGURE  
69



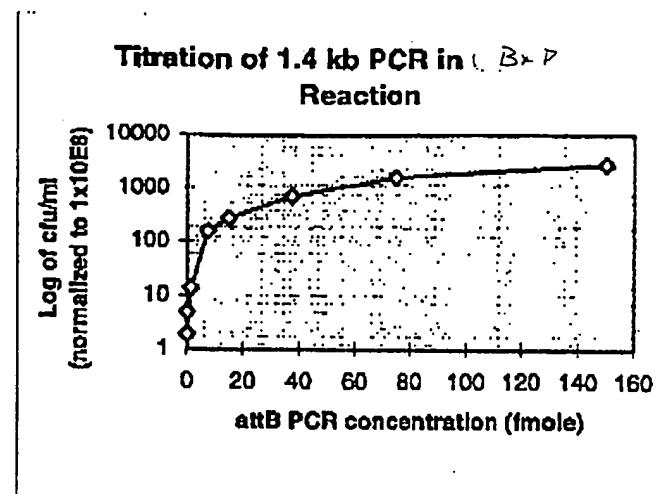
179/240  
FIGURE  
70



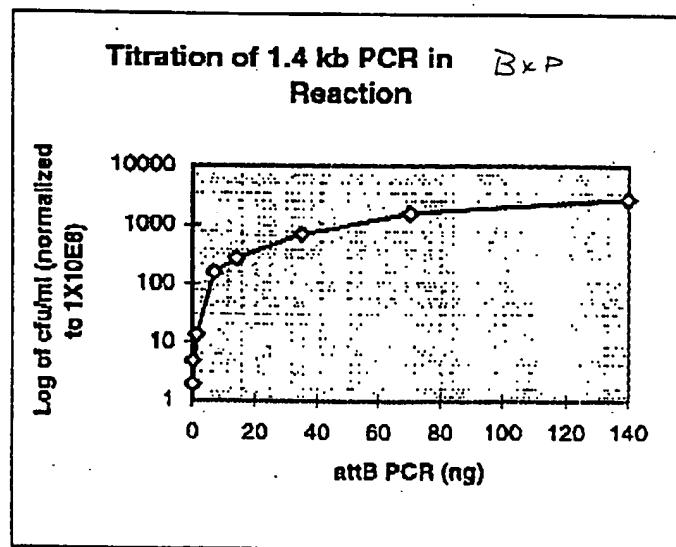
180/240

FIGURE 71

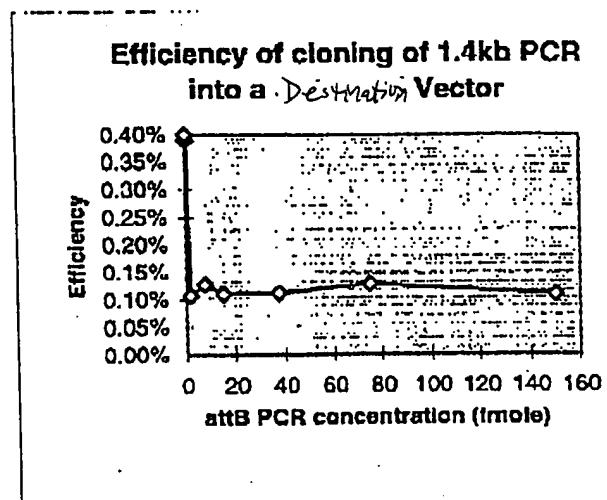
A



B



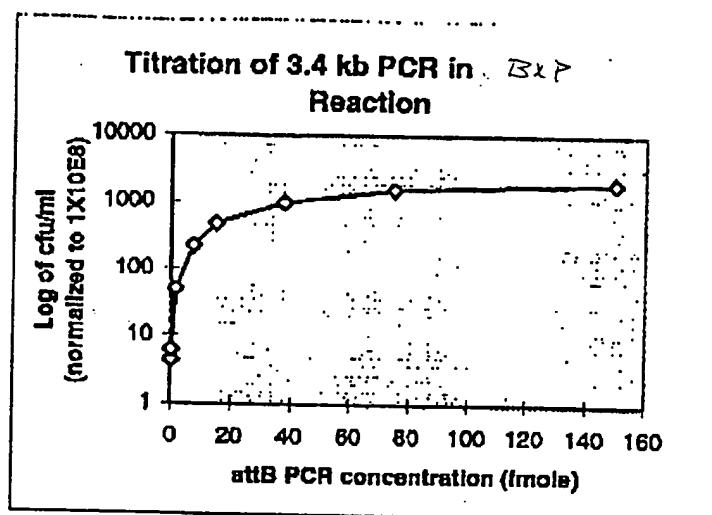
C



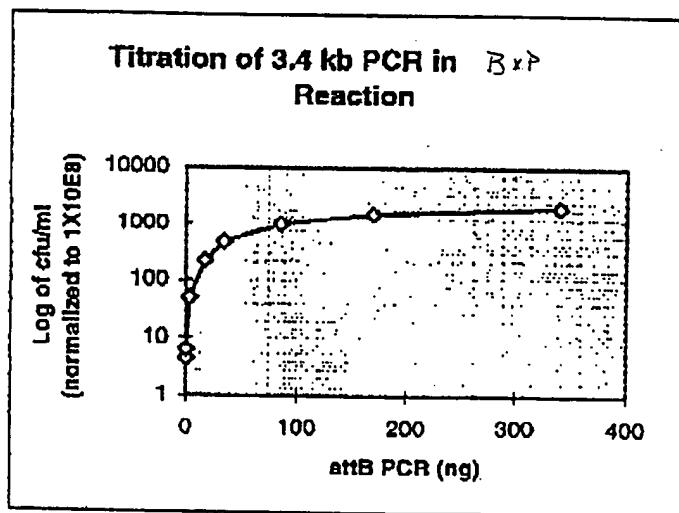
181/240

FIGURE 72

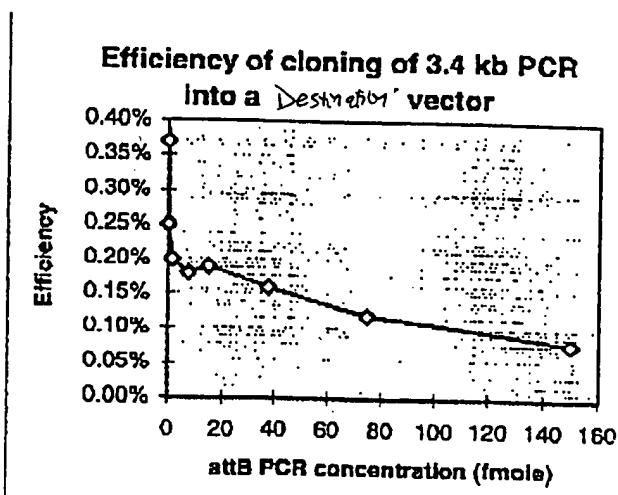
A



B



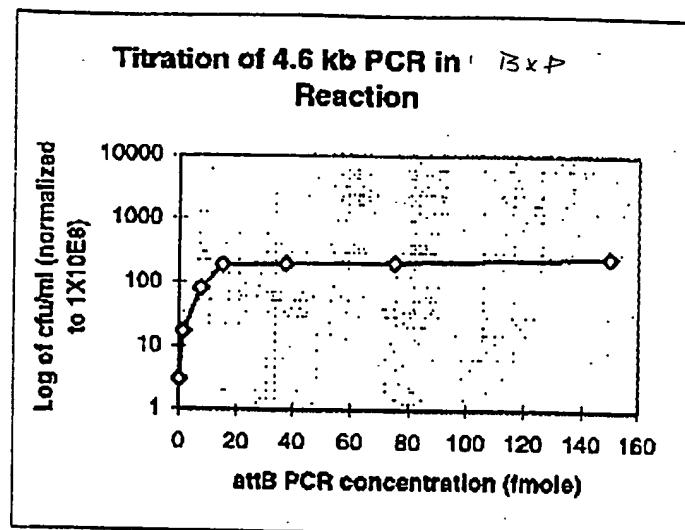
C



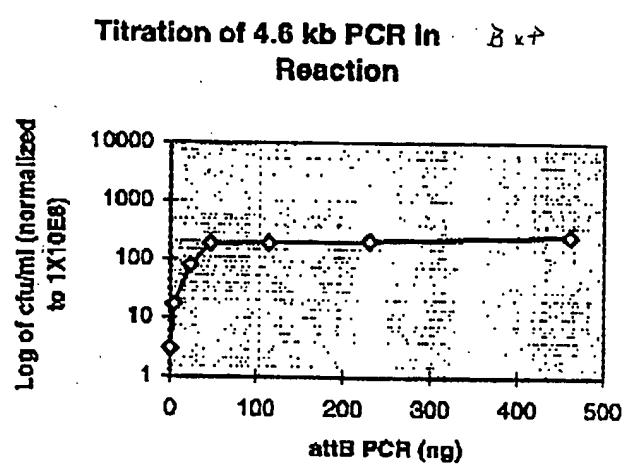
182/240

FIGURE 73

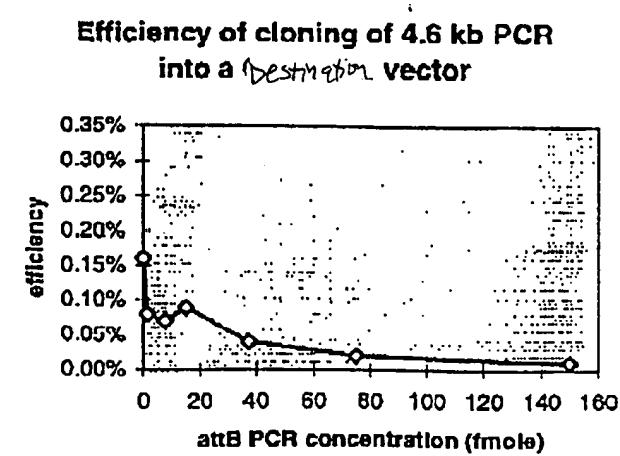
A



B



C



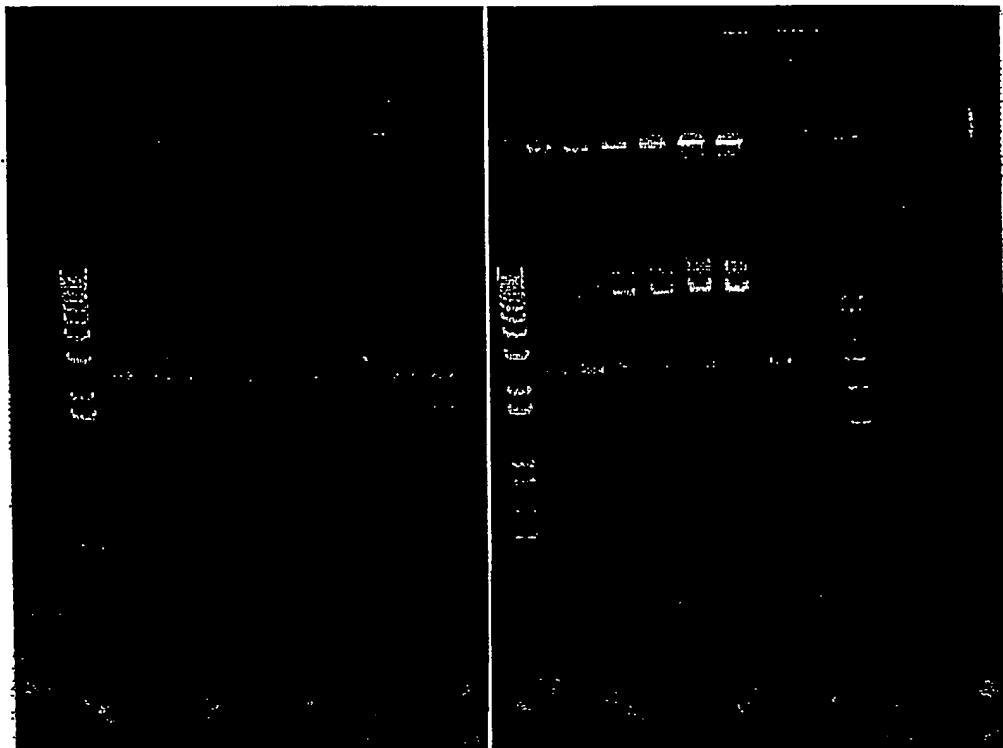
**6.9 kb PCR DNA Titration in a BxP Reaction**

FIGURE 74

184/240

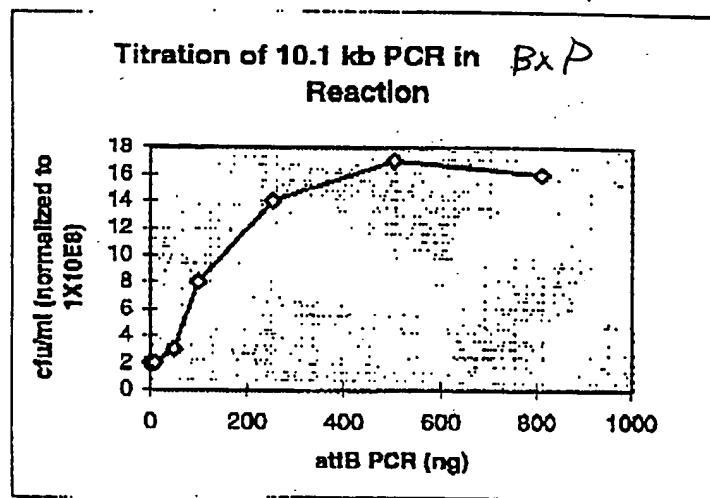


FIGURE 75-

185/260

## 10.1 kb PCR DNA Titration in $Bx^+$ Reaction

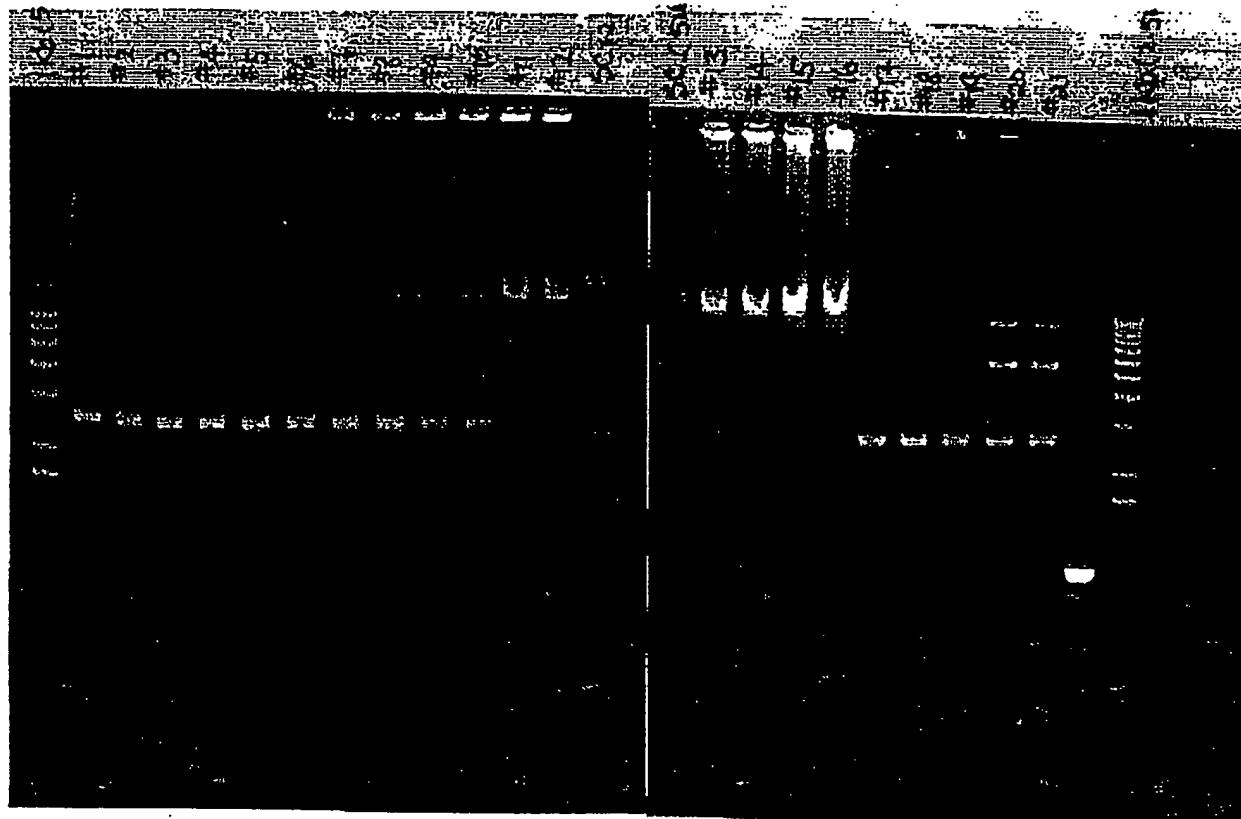


FIGURE 76

186/240

**Cloning of PCR Products of Different Sizes with the  
GATEWAY™ PCR Cloning System**

Size	fmols PCR DNA	ng PCR DNA	Cols/ml Transformation (pUC=10 <sup>8</sup> CFU/ml)	Correct Clones/Total Examined**
0.26 kb*	15	3	1223	10/10 (a)
	37.5	7.5	2815	
1.0 kb	15	10	507	49/50 (b)
	37.5	25	1447	
1.4 kb	15	14	271	48/50 (c)
	37.5	35	683	
3.4 kb	15	34	478	9/10 (a)
	37.5	85	976	
4.6 kb	15	46	190	10/10 (a)
	37.5	115	195	
6.9 kb	15	69	30 (235)**	47/50 (b)
	37.5	173	54 (463)**	

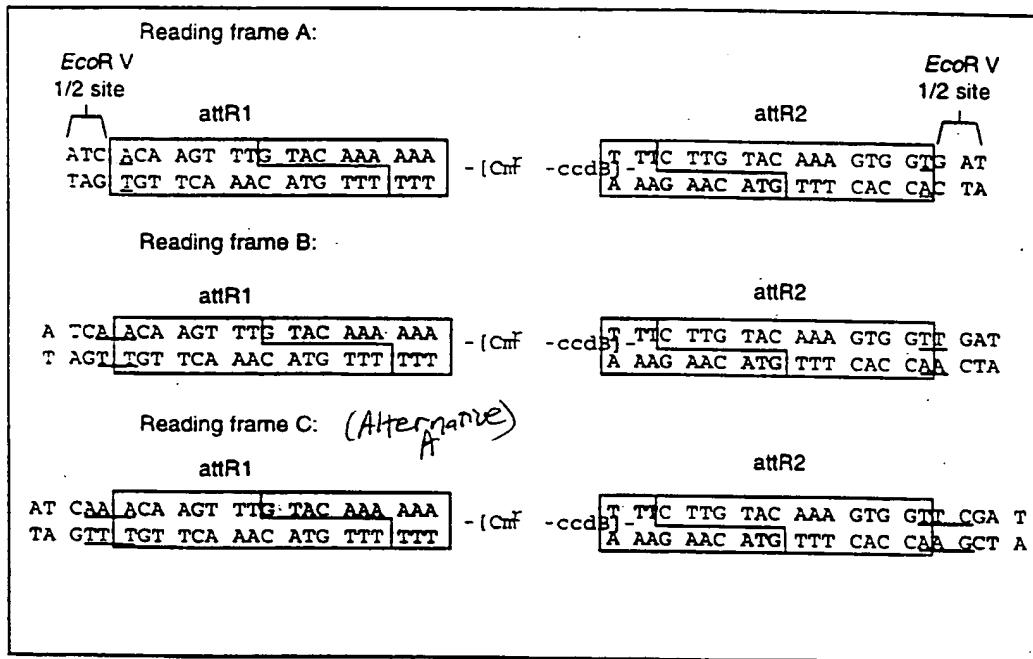
\*The 0.26 kb PCR product was used unpurified; all the others were purified by precipitation with PEG/MgCl<sub>2</sub> as described in the text of Example 9, to remove primer dimers potentially present. Standard incubations were for 60 min.

\*\*overnight incubation

- (a) DNA minipreps
- (b) ampR/kanR
- (c) tetR/kanR

**Figure 77**

187/260



**Reading frame C:** (Alternative)  
B

att R1

att R2

AT CAA ACA AGT TTE / TAC AAA AAC - [CmR-ccdB] - T TTG TAC AAA GTG GTT TGA T  
 TA GTT TGT TCA AAC ATG TTT / TTT / AAG AAC ATG TTT CAC CAA ACT A

## FIGURE 78

188/240

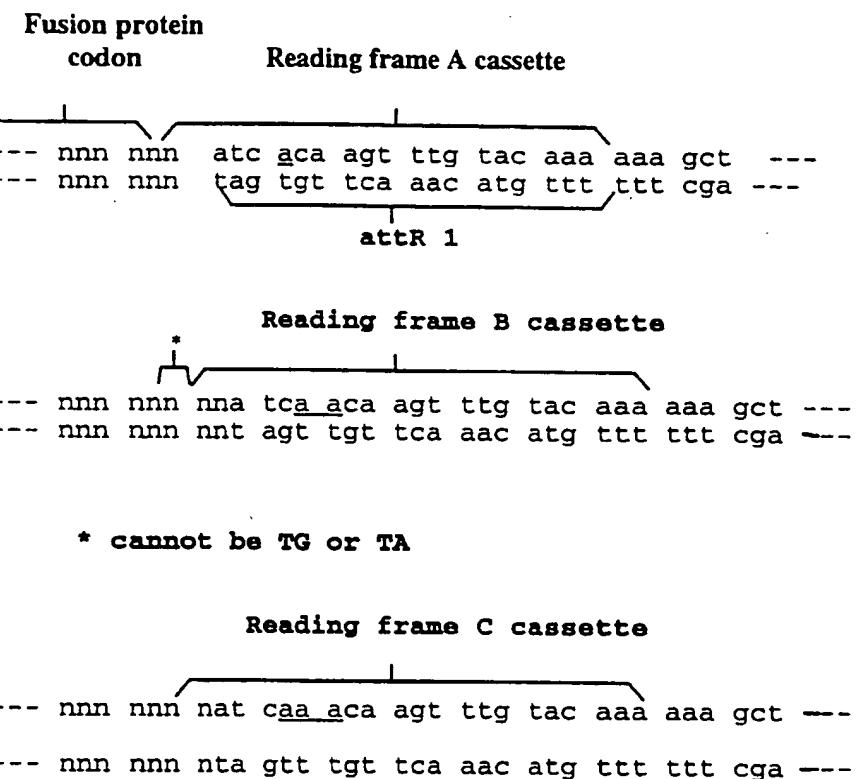
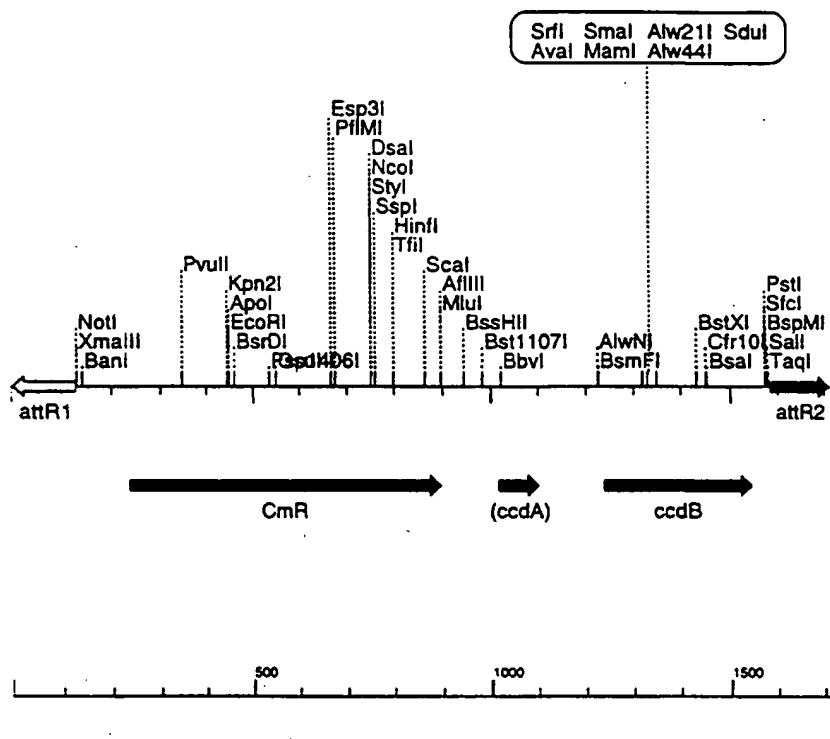


FIGURE 79

189/260



## **rfa Cassette (1711 bps)**

## FIGURE 80

190/260

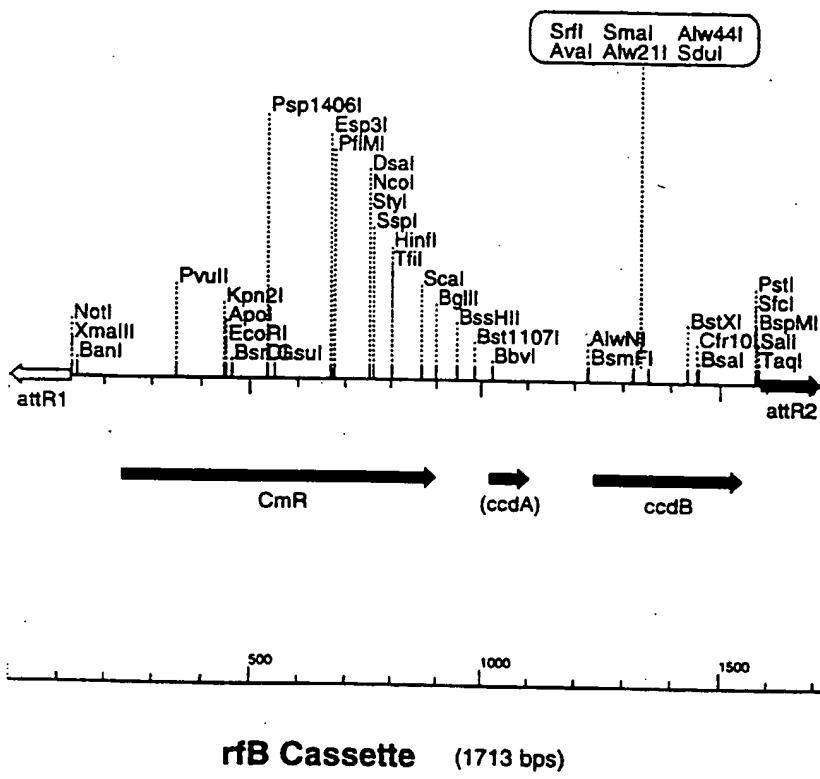
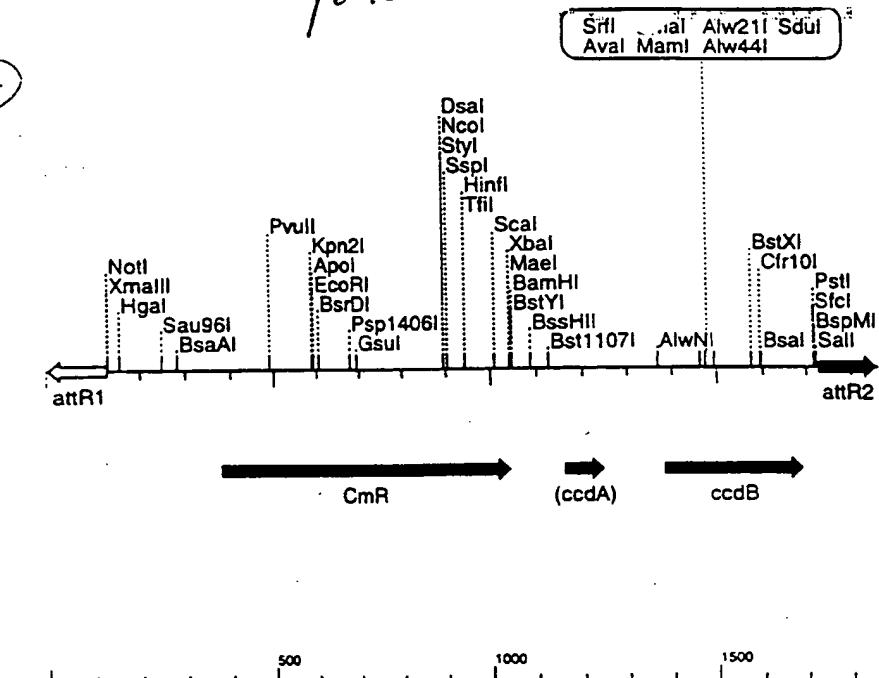
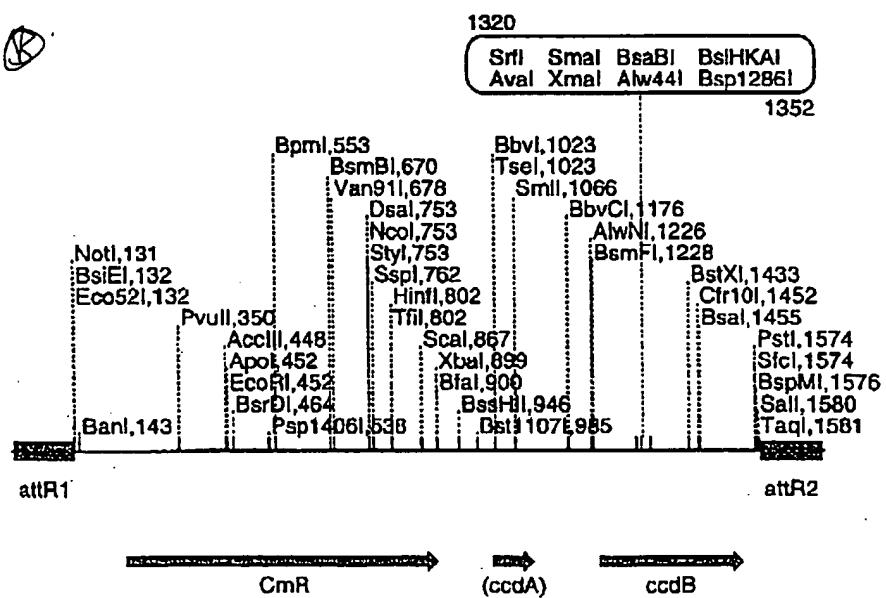


FIGURE 81



rfC Cassette (1856 bps)



rfC cassette (1715 bps)

FIGURE 82

192/240

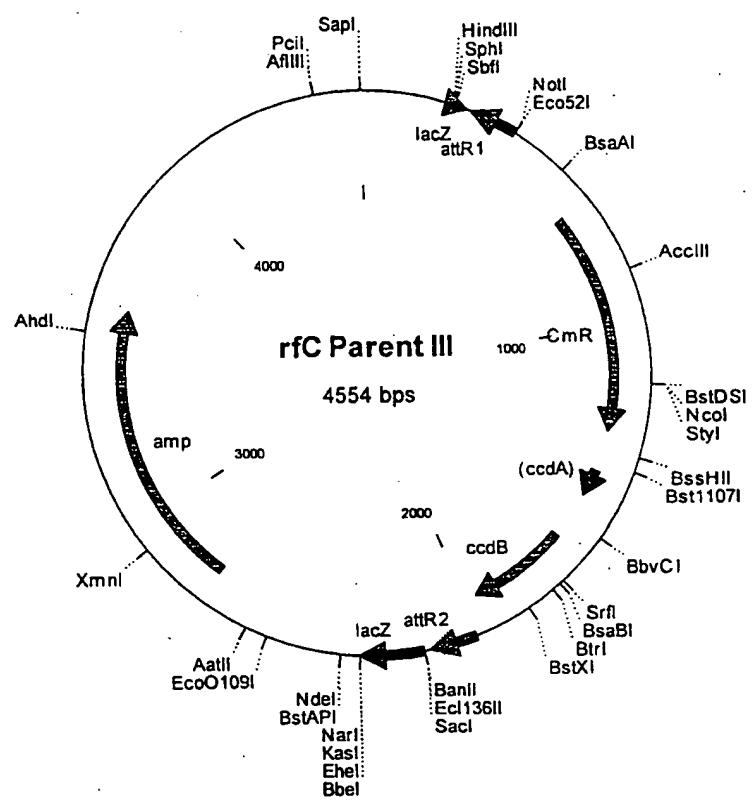


FIGURE 83 A

## prfC Parent III 4554 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
410..286	attR1
660..1319	CmR
1439..1523	inactivated ccdA
1661..1966	ccdB
2007..2131	attR2
2753..3613	amp

1 GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCG ATTCAATTAT GCAGCTGGCA  
 61 CGACAGGTTT CCCCAGCTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAAATG TGAGTTAGCT  
 121 CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT  
 181 TGTGAGCGGA TAACAATTTC ACACAGAAA CAGCTATGAC CATGATTACG CCAAGCTTGC  
 241 ATGCCTGCAG GTGCACTCTA GAGGATCCCC GGGTACCGAT ATCAAACAAG TTTGTACAAA  
 301 AAAGCTGAAC GAGAAACGTA AAATGATAA AATATCAATA TATTAATTA GATTTGCAT  
 361 AAAAACAGA CTACATAATA CTGTAAAACA CAACATATCC AGTCACTATG GCGGCCGCTA  
 421 AGTTGGCAGC ATCACCCGAC GCACTTTGCG CCGAATAAT ACCTGTGACG GAAGATCACT  
 481 TCGCAGAATA AATAAACCTT GGTGTCCCTG TTGATACCGG GAAGCCCTGG GCCAACCTTT  
 541 GGCGAAAATG AGACGTTGAT CGGCACGTA GAGGTTCCAA CTTTACCAT AATGAAATAA  
 601 GATCACTACC GGGCGTATT TTTGAGTTAT CGAGATTTTC AGGAGCTAAG GAAGCTAAAA  
 661 TGGAGAAAAA AATCACTGGA TATACCACCG TTGATATATC CCAATGGCAT CGTAAAGAAC  
 721 ATTTGAGGC ATTCAGTCA GTTGCTCAAT GTACCTATAA CCAGACCGTT CAGCTGGATA  
 781 TTACGGCTT TTTAAAGACC GTAAAGAAAA ATAAGCACAA GTTTTATCCG GCCTTTATTC  
 841 ACATTCTGC CGGCCTGATG AATGCTCATC CGGAATTCCG TATGGCAATG AAAGACGGTG  
 901 AGCTGGTGTATGGGATAGT GTTCACCCCTT GTTACACCGT TTTCCATGAG CAAACTGAAA  
 961 CGTTTCATC GCTCTGGAGT GAATACACG ACGATTTCCG GCAGTTCTA CACATATATT  
 1021 CGCAAGATGT GGCCTGTTAC GGTGAAAACC TGGCCTATT TCCCTAAAGGG TTTATTGAGA  
 1081 ATATGTTTT CGTCTCAGCC AATCCCTGGG TGAGTTTCAC CAGTTTGAT TTAAACGTGG  
 1141 CCAATATGGA CAACTTCTTC GCCCCGTTT TCACCATGGG CAAATATTAT ACGCAAGGCG  
 1201 ACAAGGTGCT GATGCCGCTG GCGATTTCAGG TTCATCATGC CGTCTGTGAT GGCTTCATG  
 1261 TCGGCAGAAT GCTTAATGAA TTACACAGT ACTGCGATGA GTGGCAGGGC GGGCGTAAT  
 1321 CTAGAGGATC CGGCTTACTA AAAGCCAGAT AACAGTATGC GTATTTGCC GCTGATTTTT  
 1381 GCGGTATAAG AATATATACT GATATGTATA CCCGAAGTAT GTCAAAAAGA GGTGTGCTAT  
 1441 GAAGCAGCGT ATTACAGTGA CAGTTGACAG CGACAGCTAT CAGTTGCTCA AGGCATATAT  
 1501 GATGTCATAA TCTCCGGTCT GGTAAGCACA ACCATGCAGA ATGAAGCCCG TCGTCGCGT  
 1561 GCCGAACGCT GGAAAGCGGA AAATCAGGAA GGGATGGCTG AGGTGCCCCG GTTTATTGAA  
 1621 ATGAACGGCT CTTTTGCTGA CGAGAACAGG GACTGGTGA ATGCAGTTA AGGTTTACAC  
 1681 CTATAAAAAGA GAGAGCCGTT ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC  
 1741 GCGGGCGA CGGATGGTGA TCCCCCTGGC CAGTCACAGT CTGCTGTCAG ATAAAGTCTC  
 1801 CCGTGAACCT TACCCGGTGG TGATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA  
 1861 TATGGCCAGT GTGCCGGTCT CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCCACCGCGA  
 1921 AAATGACATC AAAAACGCCA TTAACCTGAT GTTCTGGGG ATATAAAATGT CAGGCTCCGT  
 1981 TATACACAGC CAGTCCTGCA GTGACCTGAT GTGACTGGAT ATGTTGTTTG TTACAGTATT  
 2041 ATGTAGTCG TTTTTTATGC AAAATCTAAT TTAATATATT GATATTTATA TCATTTACG  
 2101 TTTCTCGTTC AGCTTTCTG TACAAAGTGG TTGATATCG GTACCGAGCT CGAATTCACT  
 2161 GGCGTCGTT TTACAAACGTC GTGACTGGGA AAACCCCTGGC GTTACCCAAC TTAATCGCCT  
 2221 TGCAGCACAT CCCCCCTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC  
 2281 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTG ATGCGGTATT TTCTCCTTAC  
 2341 GCATCTGTGC GGTATTTCAC ACCGCATATG GTGCACCTCTC AGTACAATCT GCTCTGATGC  
 2401 CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGGCCCT GACGGGCTTG  
 2461 TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA  
 2521 GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGACGAAAG GGCCTCGTGA TACGCCATT  
 2581 TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTCGGGG  
 2641 AAATGTGCGC GGAACCCCTA TTGTTTATT TTCTAAATA CATTCAAATA TGTATCCGCT  
 2701 CATGAGACAA TAACCCGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT  
 2761 TCAACATTTC CGTGTGCCCC TTATCCCTT TTTTGCCTC TGTGTTTGC-

FIGURE 83B

2821 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG  
2881 TTACATCGAA CTGGATCTCA ACAGCGTAA GATCCTTGAG AGTTTCGCC CGAAGAACG  
2941 TTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA  
3001 CGCCGGCAA GAGCAACTCG GTGCCGCAT ACACATTCT CAGAATGACT TGGTTGAGTA  
3061 CTCACCCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC  
3121 TGCCATAACC ATGAGTGATA ACACTGCCG CAACTTACTT CTGACAAACGA TCGGAGGACC  
3181 GAAGGAGCTA ACCGCTTTTG TGCACAAACAT GGGGGATCAT GTAACCTGCC TTGATCGTTG  
3241 GGAACCGGAG CTGAATGAAG CCATACAAA CGACGAGCGT GACACCACGA TGCCTGTAGC  
3301 AATGGCAACA ACGTTGCGCA AACTATTAAAC TGGCGAACTA CTTACTCTAG CTTCCGGCA  
3361 ACAATTAAATA GACTGGATGG AGGCGGATAA AGTTCAGGA CCACCTCTGC GCTCGCCCT  
3421 TCCGGCTGGC TGGTTTATTG CTGATAAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT  
3481 CATTGCAGCA CTGGGGCCAG ATGGTAGGC CTCCCGTATC GTAGTTATCT ACACGACGGG  
3541 GAGTCAGGCA ACTATGGAT AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT  
3601 TAAGCATTGG TAACTGTCA ACCAAGTTA CTCATATATA CTTTAGATTG ATTTAAACT  
3661 TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTT GATAATCTCA TGACCAAAAT  
3721 CCCTTAACGT GAGTTTCGT TCCACTGAGC GTCAAGACCCC GTAGAAAAGA TCAAAGGATC  
3781 TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCAACGCT  
3841 ACCAGCGGTG GTTGTGTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACTGG  
3901 CTTCAGCAGA GCGCAGATAAC CAAATACTGT CCTCTAGTG TAGCGTAGT TAGGCACCA  
3961 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAAGTGGC  
4021 TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA  
4081 TAAGGCGCAG CGGTGGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC  
4141 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCGA  
4201 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG  
4261 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTGGGGTTTC GCCACCTCTG  
4321 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGGCGG AGCCTATGGA AAAACGCCAG  
4381 CAACCGGGCC TTTTACGGT TCCCTGGCCCT TTGCTGGCCT TTTGCTACA TGTTCTTIC  
4441 TGCCTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC  
4501 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGA

FIGURE 83C

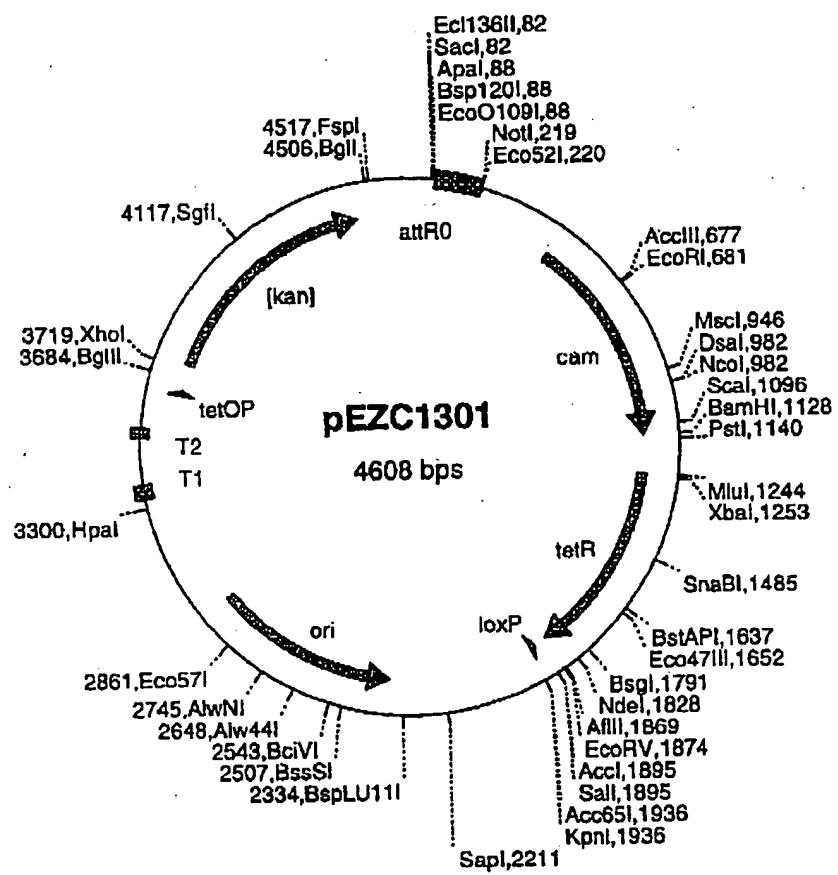


FIGURE 84

196/260

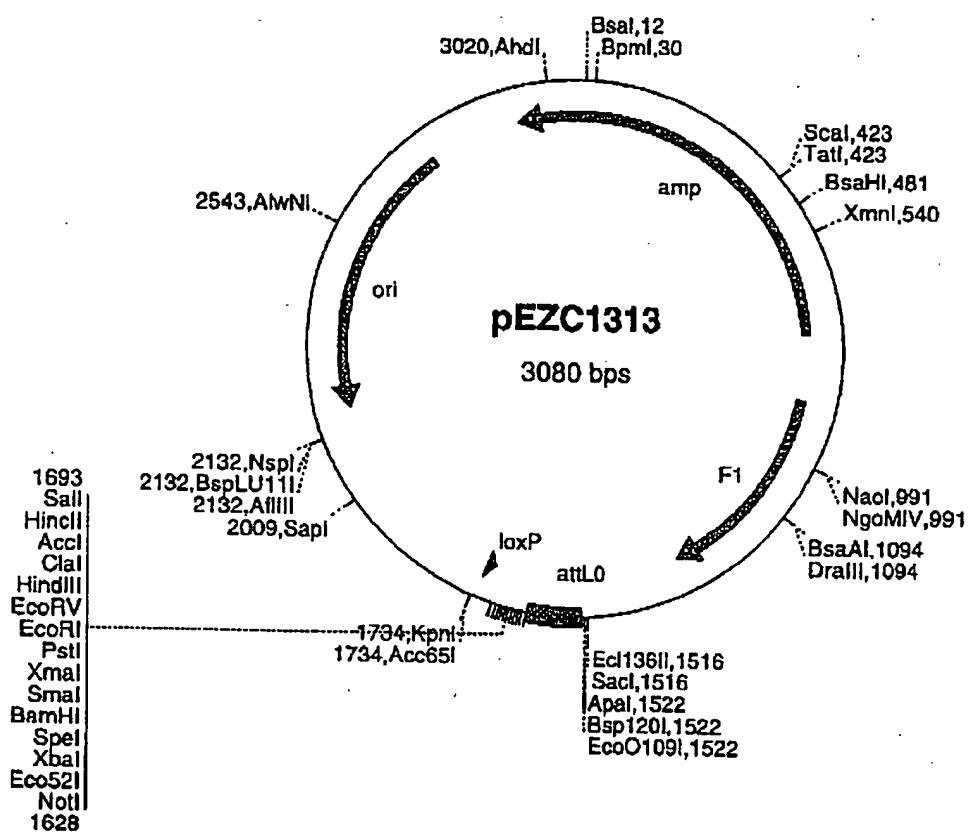


FIGURE 85

P7/240

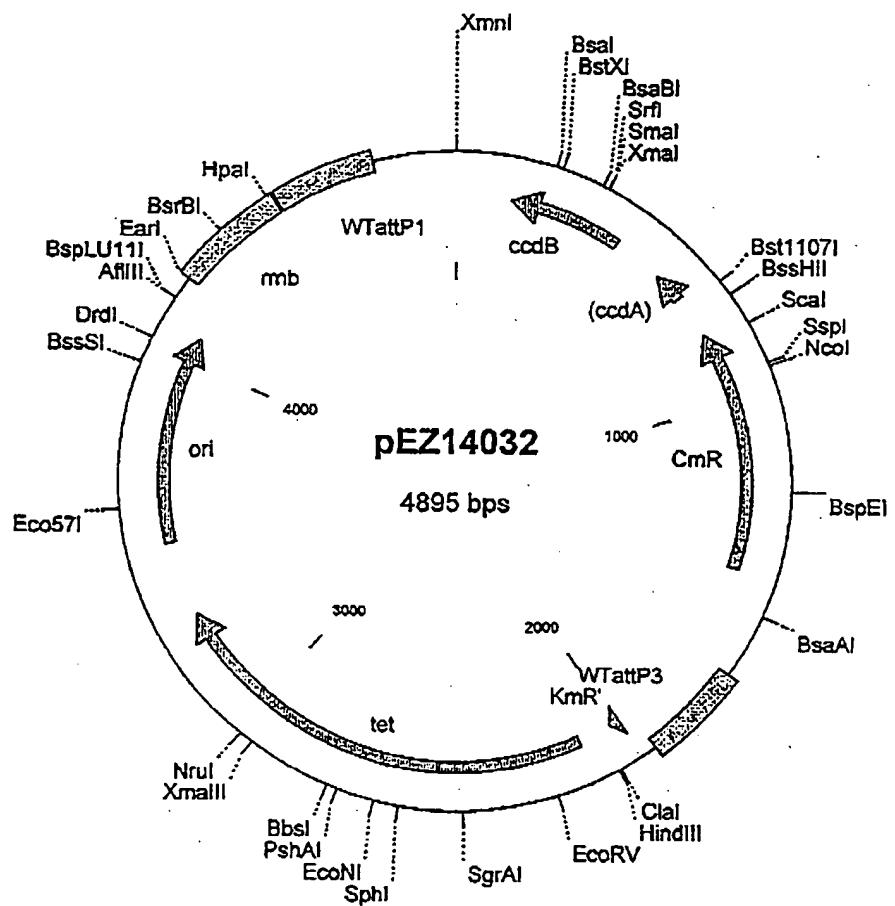


FIGURE 86

198/260

FIGURE 87

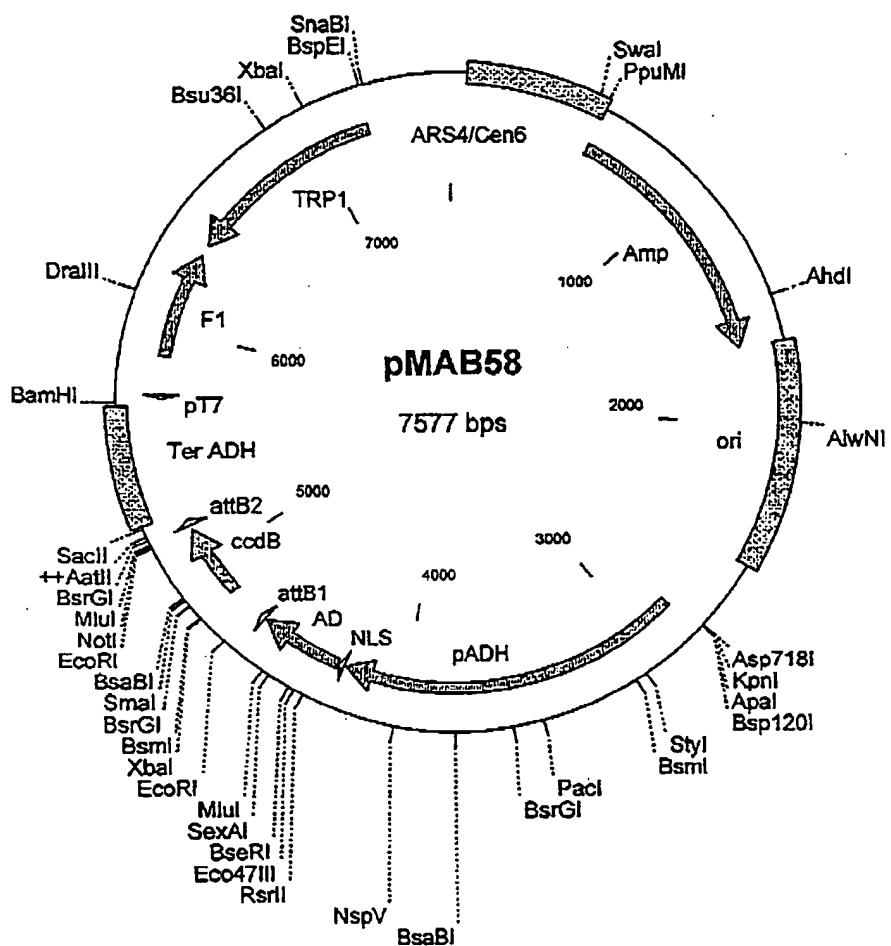
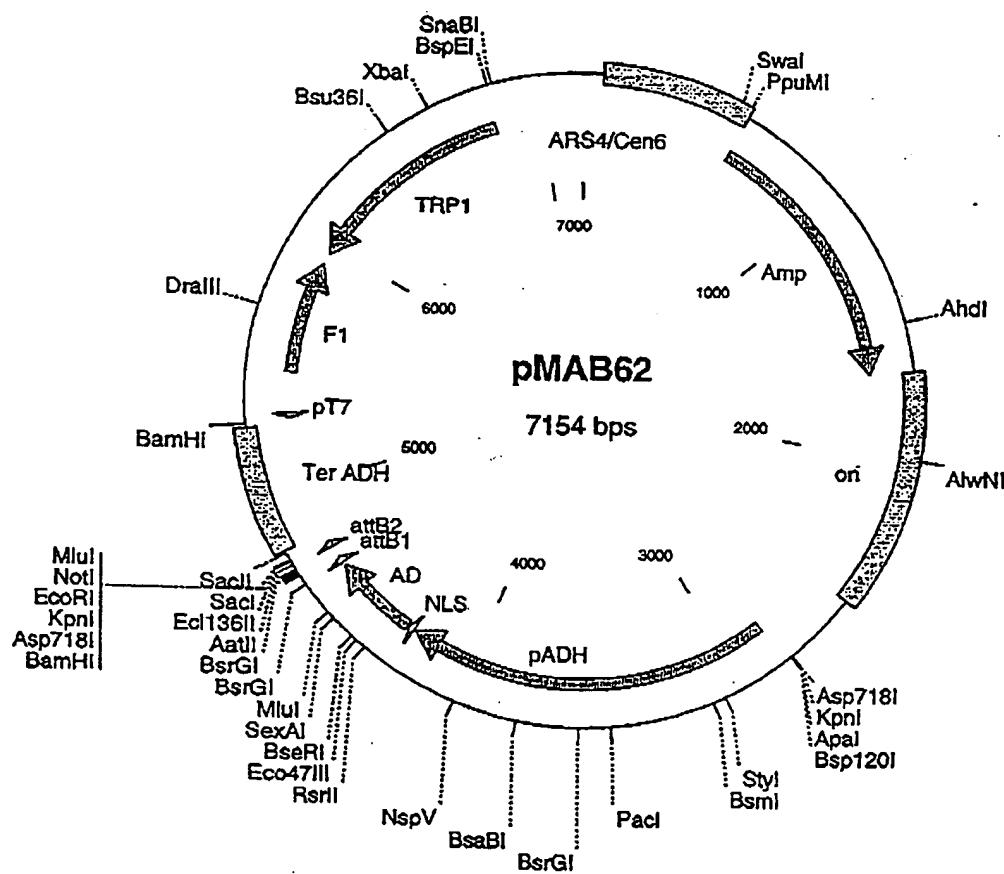
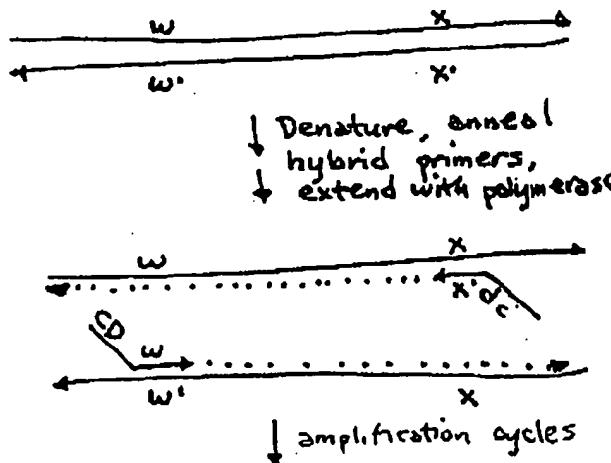


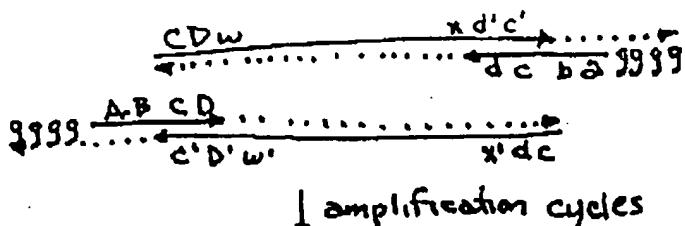
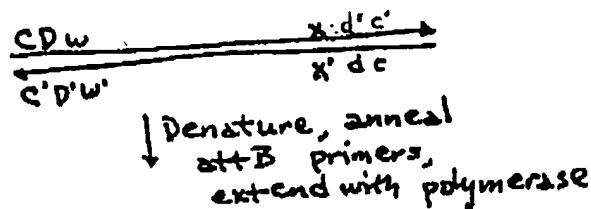
FIGURE 88



200/240

DNA to be amplified ( $5' \rightarrow 3'$ ):attB1 primer:  
gggg  $\xrightarrow{ABCD}$ attB2 primer:  
gggg  $\xrightarrow{abcd}$ 

Hybrid primers (part attB, part gene specific):

 $\underline{CD} \underline{w} \rightarrow$   
 $\underline{cd} \underline{x'} \rightarrow$ 

attB1	attB2
$\overbrace{\text{ABCD}} \text{w}$	$\overbrace{\text{x'dc'b'a'}} \text{ccaa}$
$\overbrace{\text{A'B'C'D'w'}}$	$\overbrace{\text{x'dc'b'a'}} \text{ccaa}$

FIGURE 89

201/240

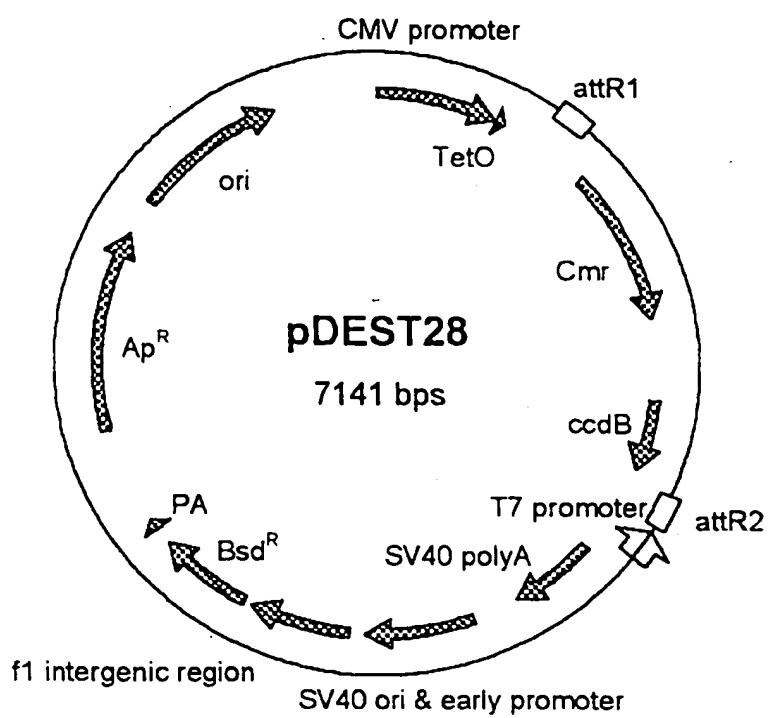


FIGURE 90A

pDEST28 7141 bp

ATGCATGTCGTTACATAACTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC  
 CGCCCATGGTACGTCATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT  
 TGACGTCAATGGTGGAGTTACGGTAAACTGCCCCTTGCCAGTACATCAAGTGTAT  
 CATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT  
 GCCCAGTACATGACCTTATGGGACTTCCTACTTGGCAGTACATCTACGTATTAGTCATC  
 GCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTGAC  
 TCACGGGGATTCCAAGTCTCCACCCCCATTGACGTCAATGGGAGTTGTTGGCACCAA  
 AATCAACGGACTTTCCAAAATGCGTAACAACCTCCGCCATTGACGCAAATGGCGGT  
 AGGCCTGACGGTGGAGGCTATATAAGCAGAGCTCCCTATCAGTGTAGAGATCTC  
 CCTATCAGTGTAGAGATCGCAGCAGCTCGTTAGTGAACCCTAGATCGCTGGAGA  
 CGCCATCCACGCTGTTGACCTCCATAGAACGACCCGGACCGATCCAGCCTCCGGACT  
 CTAGAGGATCCCTACCGGTGATATCCTCGAGCCATCAACAAAGTTGTCACAAAAAGCTG  
 AACGAGAACGTAATGATATAATCAATATATTAAATTAGATTGTCATAAAAAAAC  
 AGACTACATAATACTGAAACACAAACATATCCAGTCACTATGGCGGCCGATTAGGCAC  
 CCCAGGCTTACACTTATGCTTCCGGCTCGTATAATGTTGAGTTAGGATCC  
 GGCAGGATTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAAATCACTGGATATACCAC  
 CGTTGATATATCCCAATGGCATCGTAAGAACATTGAGGCTTCAGTCAGTTGCTCA  
 ATGTACCTATAACCAAGACCGTTCAGCTGGATATTACGGCTTTAAAGACCGTAAAGAA  
 AAATAAGCACAAGTTTATCCGGCTTATTACACATTCTGGCCGCTGATGAATGCTCA  
 TCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGTATGGGATAGTGTACCC  
 TTGTTACACCGTTTCCATGAGCAAACGTTTACACATATTCGCAAGATGTTGAGTGAATACCA  
 CGACGATTCCGGCAGTTCTACACATATTCGCAAGATGTTGAGTGAATACCA  
 CCTGGCCTATTCCCTAAAGGGTTATTGAGAATATGTTTGTCTCAGGCCAATCCCTG  
 GGTGAGTTTACCAAGTTAAACGTTGCAATATGGACAACCTCTCGCCCCCGT  
 TTTCACCATGGGCAAATATTACGCAAGGCACAAGGTGCTGATGCCGCTGGCGATTCA  
 GGTTCATCATGCCGCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACA  
 GTACTGCGATGAGTGGCAGGGGGCGTAAAGATCTGGATCCGGCTACTAAAGCCAG  
 ATAACAGTATGCGTATTGCGCGCTGATTTTGGGTATAAGAATATACTGATATGTA  
 TACCCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGAC  
 AGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCATATCTCCGGCTGGTAAGCA  
 CAACCATGCAAGAACGCCCCGCTGCTGCGTGGAAACGCTGGAAAGCGGAAATCAGG  
 AAGGGATGGCTGAGGTCGCCGGTTATTGAAATGAACGGCTTTGCTGACGAGAAC  
 GGGACTGGTGAATGCAAGTTAACGTTACACCTATAAAAGAGAGAGCCGTTACGTC  
 TTTGTGGATGTACAGAGTGTGATATTGACACGCCGGCGACGGATGGTGTACCCCTG  
 GCCAGTGCACGCTGCTGTCAGATAAAGTCTCCGTGAACCTTACCCGGTGGTGCATATC  
 GGGGATGAAAGCTGGCGCATGATGACCAACGATATGCCAGTGTGCCGGTCTCCGTTATC  
 GGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAATGACATCAAAACGCCATTACCTG  
 ATGTTCTGGGAATATAAAATGTCAGGCTCCCTATACACAGCCAGTCTGCAGGTCACCA  
 TAGTGACTGGATATGTTGTTACAGTATTATGAGTCTGTTTATGCAAATCTA  
 ATTAAATATATTGATATTATCATTTACGTTCTCGTTACGTTCTTGTCACAAAGT  
 GGTTGATGGCGGGCGCTTAGAGGGCCAAGCTTACGCGTGCATGCGACGTCAAGCTC  
 TCTCCCTATAGTGAGTCGTTACAGCTGTTACAGGCTGTTTACACGTCGACCA  
 CTGGGAAACTGCTAGCTGGATCTTGTAAGGAACCTTACCTGTTGTTGACATA  
 ATTGGACAAACTACAGAGATTAAAGCTAAGGTAATATAAAATTAAAGTGT  
 ATAATGTTAAACTAGCTGCATATGCTGCTGTTGAGAGTTTGTACTGAGTATGA  
 TTATGAAAATATTACACAGGAGCTAGTGATTCTAATTGTTGTTAGATTCA  
 CAGTCCCAGGCTCATTTCAGGCCCTCAGTCCTACAGTCTGTTACGTCAGGTCAC  
 CCATACACATTGAGGTTTACTGTTAAAAACCTCCACACCTCCCCCTGAA  
 CCTGAAACATAAAATGAATGCAATTGTTGTTAACTGTTATTGCACTTATAATGG  
 TTACAAATAAGCAATAGCATCACAAATTTCACAAATAAGCATTTTCACTGCATT  
 TAGTTGTTGCTGTCAGGCTACATCAATGTTACATGCTGGATCGATCCTGCATT  
 AATGAATCGGCCAACCGCGGGGAGAGGGCGGTTGCGTATTGCTGGCGTAATAGCGAAG  
 AGGCCCGCACCAGTCGCCCTCCAAAGCTTGCAGCCTGAATGGCGAATGGGACGCGC  
 CCTGAGCGCGCATTAAAGCGCGGGTGTGGTGGTTACGCGCAGCGTACCGCTACAC  
 TTGCCAGGCCCTAGGCCGCTCCTTCGCTTCTCCCTTCTGCCACGTTCG  
 CGGCCCTTCCCCGTCAAGCTAAATGGGGCTCCCTTACGGGTTCCGATTAGTGCTT-

FIGURE 9B

203/240

TACGGCACCTCGACCCCCAAAAAACTTGATTAGGGTGTGATGGTTACGTAGTGGGCCATCGC  
 CCTGATAGACGGTTTTCGCCCTTGACGTTGGACTCCACGTTCTTAATAGTGGACTCT  
 TGTTCAAAACCTGAAACAACACTCAACCTATCTCGGTCTATTCTTTGATTTATAAGGGA  
 TTTGCCGATTCGCCATTGGTAAAAATGAGCTGATTAAACAAATATTAACGCAT  
 ATTTAACAAAATATTAACGTTACAATTCCGCTGATGCCGTATTTCTCCTACGCAT  
 CTGTCGGTATTCACACCGCATACCGGATCTGCGCAGCACCATGCCGTGAAATAACCT  
 CTGAAAGAGGAACCTGGTTAGGTACCTCTGAGGCGGAAAGAACCGCTGTGGAATGTGT  
 GTCAGTTAGGGTGTGAAAGTCCCCAGGCTCCCCAGCAGGAGAAGTATGCAAAGCATGC  
 ATCTCAATTAGTCAGCAACAGGTGGAAAGTCCCCAGGCTCCCCAGCAGGAGAAGTA  
 TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCCCTAACTCCGCCATCC  
 CGCCCCCTAACCTCCGCCAGTCCGCCATTCTCCGCCATGGCTGACTAATTTTTA  
 TTTATGCAGAGGCCAGGCCGCTGCCCTGAGCTATTCCAGAAGTAGTGAGGAGGCT  
 TTTTGAGGCCCTAGGCTTGCACAAAGCTGATTCTCTGACACAAACAGTCTCGAACT  
 TAAGACCATGGCCAAGCCTTGTCTCAAGAAGAACATCCACCTCATTGAAAGAGCAACGGC  
 TACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGGCAGCTCTCTAG  
 CGACGGCGCATCTCACTGGTGTCAATGTATATCATTACTGGGGACCTGTGAGA  
 ACTCGTGGTGTGGCACTGCTGCTGCCAGCTGGCACCTGACTTGATCGTGC  
 GATCGGAAATGAGAACAGGGCATCTGAGCCCCCTGCCAGGGTGGCACAGGTGCTTCT  
 CGATCTGCATCCTGGATCAAAGCCATAGTGAAGGACAGTGATGGACAGCGACGGCAGT  
 TGGGATTGCTGAATTGCTGCCCTCTGGTTATGTGTGGAGGGCTAACGACTCGTGGCCG  
 AGTTGAAATGACCGACCAAGCGACGCCAACCTGCCATACGATGCCGCAATAAAATA  
 TCTTATTTCTTACATCTGTGTGGTTTTGTGTGAATCGATAGCGATAAGGATC  
 CGCGTATGGTGCACTCTCACTACAATCTGCTCTGATGCCCATAGTTAACCGAGCCCCGA  
 CACCGCCAACACCCGCTGACGCCCTGACGGGCTGCTGCTCCGGCATCGCTTAC  
 AGACAAAGCTGTGACCGTCTCCGGAGCTGCATGTCAGAGGTTTCAACGTGACCG  
 AAACCGCGAGACGAAAGGGCTCGTGTACGCCATTAGGTTAATGTCATGATA  
 ATAATGGTTCTTAGACGTCAGGTGGCACTTTGGAAATGTGCGCGAACCCCTATT  
 TGTTTATTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCGTATAA  
 ATGCTTCAATAATATTGAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCT  
 ATTCCCTTTTGCAGGCCATTGCTTCTGTTTGCTCACCCAGAAACGCTGGTGA  
 GTAAAAGATGCTGAAGATCAGTGGGTGCACGAGTGGTTACATCGAACTGGATCTCAAC  
 AGCGGTAAGATCCTGAGAGTTTGCAGGCCATTGCTGCTGGGAAAGAACGTTTCAATGAGC  
 ACTTAAAGTTCTGCTATGTGGCGGGTATTATCCGTTATGACGCCGGCAAGAGCAACTCGG  
 CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCAT  
 CTTACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAAC  
 ACTGCGGCAACTTACTTCTGACAACGATCGGAGGACGAAGGAGCTAACCGCTTTTG  
 CACAACATGGGGATCATGTAACCTGCCCTGATCGTGGGAACCGGAGCTGAATGAAGCC  
 ATACCAAACGAGCGTGACACCACGATGCCGTAGCAATGCCAACAGTTGCGCAA  
 CTATTAACTGGCGAACTACTACTCTAGCTCCCGCAACAATTAAATAGACTGGATGGAG  
 GCGGATAAGTTGCAAGGACCACTCTGCGCTGGCCCTTCCGGCTGGCTGGTTATTGCT  
 GATAAACTGGAGCCGTGAGCGTGGGTCTCGCGGTATCATTGACGACTGGGGCCAGAT  
 GGTAAGCCCTCCGTATCGTAGTTATCTACAGACGGGAGTCAGGCAACTATGGATGAA  
 CGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAGCATTGTAACTGTCAGAC  
 CAAGTTACTCATATACTTTAGATTGATTTAAACTTCATTAAATTAAAGGATC  
 TAGGTGAAGATCCTTTGATAATCTCATGACAAAATCCCTAACGTGAGTTTGTTC  
 CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTCTG  
 CGCGTAATCTGCTGCTTGCACACAAAAACCCACCGCTACAGCGGTGGTTTGTGCG  
 GATCAAGAGCTACCAACTCTTTCGAGGTAACCTGGCTCAGCAGAGCGCAGATACCA  
 AATACTGCTCTCTAGTGTAGCCGTAGTTAGGCCACCTCAAGAAACTCTGAGCACCG  
 CCTACATACCTCGCTCTGTAATCCGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCG  
 TGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGA  
 ACGGGGGGTTCTGACACACAGCCCAGCTGGAGCGAACGACCTACACCGAAGTAC  
 CTACAGCGTGAGCATTGAGAAAGCGCCACGCTCCGAAGGGAGAAAGGCGGACAGGTAT  
 CCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGAAACGCC  
 TGGTATCTTATAGTCCTGTCGGGTTTGCACCTCTGACTTGAGCGTCGATTTGTGA  
 TGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCCGCTTTTACGGTTC  
 CTGGCCTTTGCTGCCCTTGTCACTGTTCTGCTTACATGTTCTTCTGCTTATCCCCTGATTCTGTG  
 GATAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTCGCCAGCCGAACGACCGAG-

FIGURE 90C

20h/24h

CGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCCTCTCCCC  
GCGCGTTGGCCGATTCAATTAATGCAGAGCTTGCATTCGCGCGTTTCAATATTATTGA  
AGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAAT  
AAACAAATAGGGTTCCGCCACATTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACC  
ATTATTATCATGACATTAACCTATAAAAATAGGCGTAGTACGAGGCCCTTCACTCATTA  
G

FIGURE 9D

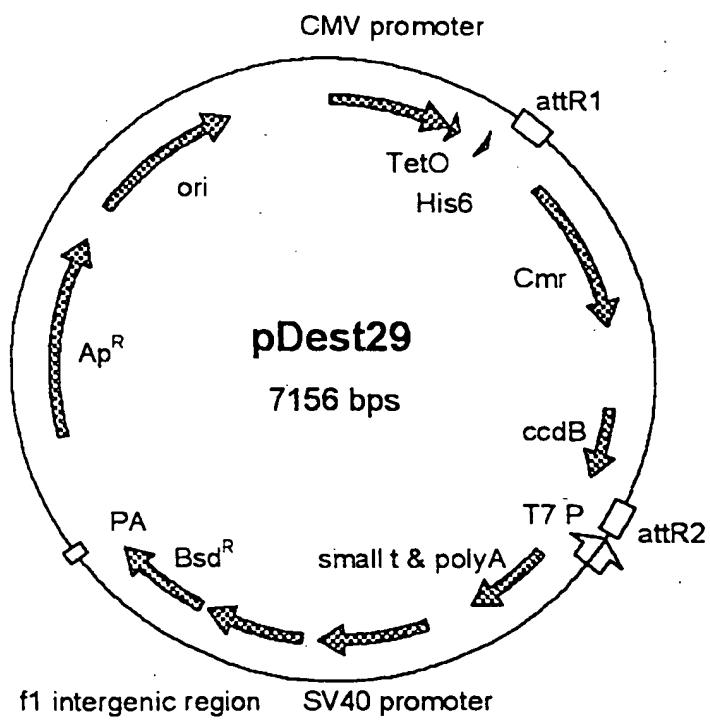


FIGURE 91 A

pDEST29 7156 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC  
 CGCCCATTGACGTCATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT  
 TGACGTCATAATGGGTGGAGTATTCACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT  
 CATATGCCAAGTACGCCCTATTGACGTCATAACGGTAAATGGCCCGCCTGGCATTAT  
 GCCCAGTACATGACCTTATGGGACTTCCACTTGGCAGTACATCAACGGCTGGATAGCGGTTGAC  
 GCTATTACCATGGTGTGGAGTACATCAACGGCTGGATAGCGGTTGAC  
 TCACGGGATTCCAAGTCTCCACCCATTGACGTCATAAGGAGTTGGCACC  
 AATCAACGGGACTTCCAAAATGTCGTAACAACCTCCGCCATTGACGCAAATGGCGGT  
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCCCTATCAGTGTAGAGATCTC  
 CCTATCAGTGTAGAGATCGTCAGCAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGA  
 CGCCATCCACGCTGTTGACCTCCATAGAACACACCAGCAGCTCCAGCCTCCGGACC  
 ATGGCGTACTACCACCATCACCATCACACCGGTGATATCCTCGAGCCCACACAGT  
 TTGTACAAAAAAGCTGAACGAGAACGTAATGATATAAATATCAATATATTAAATTAG  
 ATTTTGATACAAAACAGACTACATAACTGTAAACACAATATCCAGTCACTATGG  
 CGGCCGATTAGGCACCCAGGTTACACTTATGCTCCGGCTCGTATAATGTGTGGA  
 TTTGAGTTAGGATCCGGCAGATTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAA  
 TCACTGGATATACCACCGTTGATATATCCAAATGGCATCGTAAAGAACATTGAGGCAT  
 TTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAAGCTGGATATTACGGCCTTT  
 TAAAGACCGTAAAGAAAATAAGCACAAGTTTATCCGGCTTATTACATTCTGCC  
 GCCTGATGAATGCTCATCGGAATTCCGTATGGCAATGAAAGACGGTAGCTGGTGTAT  
 GGGATAGTGTTCACCCCTGTTACACCGTTTCCATGAGCAAACGTTTATCGC  
 TCTGGAGTGAATACCACGACGATTCCGGCAGTTCTACACATATATTGCAAGATGTGG  
 CGTGTACGGTAAAACCTGGCTATTCCCTAAAGGTTATTGAGAATATGTTTCG  
 TCTCAGCCAATCCCTGGGTGAGTTTACCAAGTTAACGTGGCCAATATGGACA  
 ACTTCTCGCCCCGTTTACCATGGCAAATATTACGCAAGGGCACAAGGTGCTGA  
 TGCCGCTGGCGATTAGGTTACATGCCGTCTGTGATGGCTTCCATGTCGGCAGAACATGC  
 TTAATGAATTACAACAGTACTCGGATGGCAGGGCAGGGCGTAAACGCGTGGATCCG  
 GCTTACTAAAAGCCAGATAACAGTATGCTATTGCGCCTGATTITGCGGTATAAGAA  
 TATATACTGATATGTATACCGAAGTATGTCAAAAGAGGGTGTGCTATGAAAGCAGCGTAT  
 TACAGTGCAGTTGACAGCGACAGCTATCAGTGTCTCAAGGCATATATGATGTCAATATC  
 TCCGGTCTGGTAAGCACAACCATGCGAAATGAAGCCGTCGCTCGGTGCCAACGCTGG  
 AAAGCGGAAATCAGGAAGGGATGGCTGAGGTGCCCCGGTTATTGAAATGAACGGCTCT  
 TTTGCTGACGAGAACAGGGACTGGTGAATGCAAGTTAACGGTTACACCTATAAAAGAGA  
 GAGCCGTTATCGCTGTTGAGTGTACAGAGTGTATTGACACGCCGGGGCGACG  
 GATGGTGTACCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAACGCTCCGTGAACCTTA  
 CCGGTGGTGCATATCGGGGATGAAAGCTGGCCTGACGCCAGTATGGCAGTGT  
 GCCGGTCTCCGTATCGGGGAGAAGTGGCTGATCTCAGGCCACCGCAAATGACATCAA  
 AAACGCCATTAACTGTGTTCTGGGAAATATAATGTCAGGCTCCGTATACACAGCCA  
 GTCTGCAGGTGACCATAGTGTACTGGGATATGTTGTTACAGTATTATGTTAGTCTGTT  
 TTTTATGAAAACTAAATTAAATATTGATATTATACATTACGTTCTCGTTCA  
 CTTTCTGTACAAAGTGGTGTGGCTGAGGGCCCAAGCTTACCGTGCAT  
 GCGACGTCATAGCTCTCCCTATAGTGTGAGTGTGTTACAGTACTGGCCTGGCTCGT  
 TTTACACGTCGTGACTGGAAAATGCTAGCTGGGATCTTGTGAAGGAACCTTACTT  
 CTGTTGTTGACATAATTGGACAAACTACACTACAGAGATTAAAGCTCAAGGTAATAT  
 AAAATTAAAGTGTATAATGTGTTAAACTAGCTGCATATGCTGCTGTTGAGAGTTT  
 GCTTACTGAGTGTGTTACAGTCCCAAGGCTCATTCAGGCCCTCAGTCTCACAGTCTGTT  
 TGTATTAGATTACAGTCCCAAGGCTCATTCAGGCCCTCAGTCTCACAGTCTGTT  
 CATGATCATAATCAGCCATACCAACATTGTTAGAGGTTACTGCTTAAAAAACCTCCC  
 ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTAACTGTTAT  
 TGCACTTATAATGGTACAAATAAGCAATAGCATCACAAATTTCACAAATAAGCATT  
 TTTTCACTGCATTCTAGTTGTTGTCACACTCATCAATGTATCTTATCATGCTG  
 GATCGATCCTGCAATTGAATGGCCAACGCCGGAGAGGGCGTTGCTATTGGCT  
 GCGTAATAGCGAAGAGGCCCGCACCAGTCGCCCTCCAAACAGTGTGCGCAGCCTGAATG  
 GCGAATGGGACGCCCTGTAGCGCGCATTAGCGCAGGGGTGTTGTTACGCGCA  
 GCGTGAACGCTACACTGCCAGCGCCCTAGGCCGCTCCTTCGCTTCTCCCTCCT  
 TTCTGCCACGTTGCCGCTTCCCGTCAAGCTCTAAATGGGGCTCCCTTGGGT-

FIGURE 91B

TCCGATTAGTGCCTTACGGCACCTCGACCCCCAAAAACTGATTAGGGTATGGTCAC  
 GTAGTGGGCCATCGCCCTGATAGACGGTTTCGCCCTTGACGGAGTCCACGTTCT  
 TTAATAGTGGACTCTGTTCAAAGTGAACAACACTCAACCCATCTCGGTCTATTCTT  
 TTGATTATAAGGGATTTCGCCATTGGTAAAGGAGCTGAGCTGATTTAAC  
 AAATATTAAACGCGAATTAAACAAATATTAACTGTTACAATTGCCCTGATGCGGTAT  
 TTTCTCCTACGCATCTGCGGTATTACACCGCATACGGATCTGCGCAGCACCAT  
 GCCCTGAAATAACCTCTGAAAGAGGAACCTGGTAGGTACCTCTGAGGCGGAAAGAAC  
 AGCTGTGAATGTGTGTCAGTTAGGGTGGAAAGTCCCCAGGCTCCCCAGCAGGAGAA  
 GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC  
 CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATACTCCGCC  
 TAACCTCCGCCATCCGCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCATGGCT  
 GACTAATTTTTATTTATGAGAGGCCAGGGCCCTCGGCTTGAGCTATTCCAGA  
 AGTAGTGGAGGAGGCTTTTGGAGGCCCTAGGTTGCAAAAGCTGATTCTGACA  
 CAACAGTCTGAACCTAACGACCATGGCAAGCCTTGTCTAAGAAGAATCCACCCCTCAT  
 TGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTGAAAGACTACAGCGTCGCCAG  
 CGCAGCTCTCTCTAGCGACGGCCGATCTCACTGGTAGCTAATGTATATCATTAACTGG  
 GGGACCTTGTGAGAAGCTGGTAGCTGGGACTGCTGCTGCGCAGCTGGCACCT  
 GACTTGTATCGTCGCGATCGGAAATGAGAACAGGGCATCTTGAGCCCTGCGGACGGTG  
 CCGACAGGTGCTTCGATCTGATCTGGATCAAAGCCATAGTGAAGGACAGTGTGAG  
 ACAGCCGACGGCAGTTGGGATTCTGAAATTGCTGCTGCCCTTGTTATGTGTGGGAGGGCTA  
 AGCACTCGTGGCCGAGTTGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAT  
 GGCCGCAATAAAATCTTATTTCATTACATCTGTTGGTTTTGTGAAATCG  
 ATAGCGATAAGGATCCGCTATGGTGACTCTCAGTACAATCTGCTCTGATGCCCATAG  
 TTAAGCCAGCCCCACACCGCCAACACCGCTGACGCCCTGACGGGCTTGTCTGCTC  
 CCGCATCCGCTTACAGACAAGCTGTGACCGCTCCGGAGCTGATGTGCAAGAGGTT  
 TCACCGTCATCACCAGCGCGAGACGAAAGGGCTCGTGTACGCCATTTCAG  
 GTTAATGTCATGATAATAATGGTTCTAGACGTCAGGTGGCAGTTCTGGGAAATGTG  
 CGCGGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCCGCTCATGAGA  
 CAATAACCCCTGATAAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGTATTCAACAT  
 TTCCGTGTCGCCCTTATTCCCTTTTGCGGCTTGCCTCTGTTGCTCACCA  
 GAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGTGACAGAGTGGTACATC  
 GAACTGGATCTCAACAGCGTAAGATCCTTGAGAGTTCTGCCCGAAGAACGTTTCCA  
 ATGATGAGCACTTTAAAGTCTGCTATGTGGCGGGTATTATCCGTATTGACGCC  
 CAAGAGCAACTCGTCGCCCATACACTATTCTCAGAATGACTGGTGAGTACTCACCA  
 GTCACAGAAAAGCATCTAACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCC  
 ACCATGAGTGTAAACACTGCCAACCTTACTCTGACAACGATCGGAGGACCGAAGGAG  
 CTAACCGTTTGCACAAACATGGGGATCATGTAACTGCCCTGATGTTGGGAAACCG  
 GAGCTGAATGAAGCCATACAAACGAGCGTGACACCACGATGCTGTAGCAATGGCA  
 ACAACGTTGCGCAAACATTAACGCGAACTACTTACTCTAGCTCCCGCAACAATT  
 ATAGACTGGATGGAGGCGATAAAAGTTGCAAGGACCACTCTGCGCTCGGCCCTCCGGCT  
 GGCTGGTTATTGCTGATAAAATCTGGAGCCGGTGAGCTGGGCTCGCGGTATCTGCA  
 GCACTGGGCCAGATGGTAAGCCCTCCGTATCGTAGTTATCTACACGACGGGAGTCAG  
 GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAGCAT  
 TGGTAACTGTCAGACCAAGTTACTCATATACTTTAGATTGATTAAAACCTCATT  
 TAATTAAAAGGATCTAGGTGAAGATCCTTTGATAATCTCATGACCAAAATCCCTAA  
 CGTAGTTCTGTTCACTGAGCGTCAAGACCCGTAGAAAAGATCAAAGGATCTCTTGA  
 GATCCTTTCTGCGCTAATCTGCTGTTGCAAAACAAAAAACCGCTACAGCG  
 GTGGTTTGTGCGGATCAAGAGCTACCAACTTTTCCGAAGGTAACTGGCTTCAGC  
 AGAGCGCAGATACCAAATACTGCTCTAGTGAGCCAGTTAGGCCACCACTTCAAG  
 AACTCTGTAGCACCGCCTACATACCTCGCTGCTAATCTGTTACCGACTGGCTGCTGCC  
 AGTGGCGATAAGCTGCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAAGGCG  
 CAGCGGTGGCTGAACGGGGGTTCTGACACACAGCCCAGCTGGAGCGAACGACCTAC  
 ACCGAACGTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTCCGAAGGGAGA  
 AAGGCGGACAGGTATCCGTAAGCGCAGGGTGGAAACAGGAGAGCGCACGAGGGAGCTT  
 CCAGGGGAAACGCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAG  
 CGTCGATTTTGATGCTGTCAGGGGGCGAGCCTATGAAAACGCCAGCAACGCG  
 GCCTTTTACGGTCTGCCCTTGTGGCTTGTGCTCACATGTTCTTCTGCGTTA  
 TCCCTGATTCTGTTGATAACCGTATTACCGCTTGTGAGTGTGATACCGCTGCCCG-

208/240

AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGGCCAATACGC  
AAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAAATGCAGAGCTTGCATTCGCGCGTT  
TTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATAACATATTGAA  
TGTATTAGAAAATAAACAAATAGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCT  
GACGTCTAAGAAACCATTATTATCATGACATTAAACCTATAAAAATAGGCGTAGTACGAGG  
CCCTTCACTCATTAG

FIGURE 91D

209/240

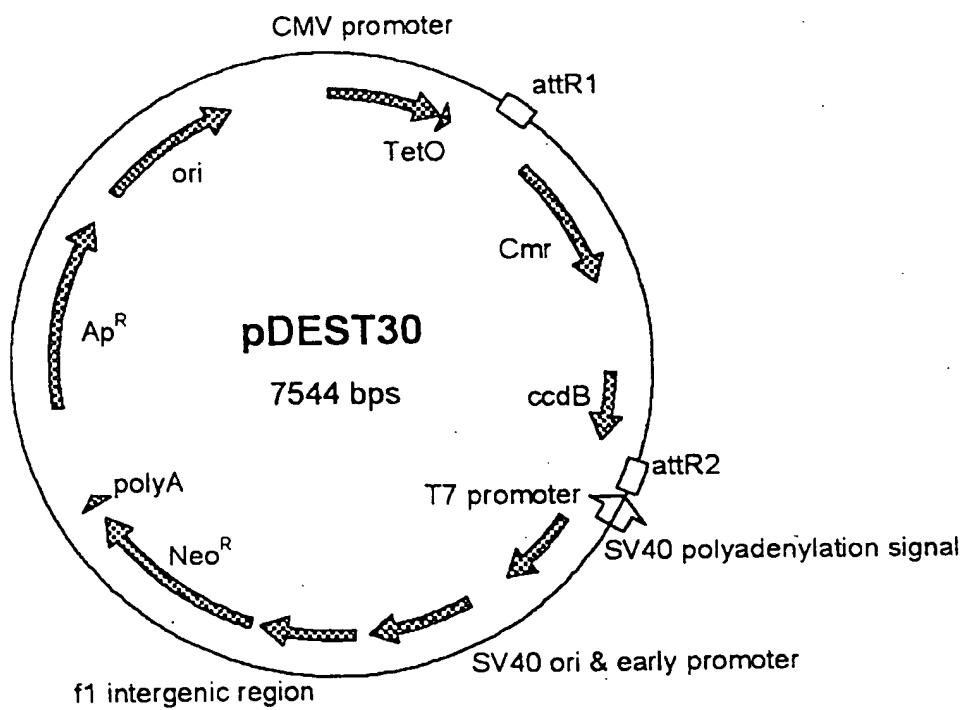


FIGURE 92A

pDEST30 7544 bp

ATGCATGTCGTTACATAACTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC  
 CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT  
 TGACGTCAATGGGTGGAGTATTCACGGTAAACTGCCACTTGGCAGTACATCAAGTGTAT  
 CATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT  
 GCCCAGTACATGACCTATGGGACTTCTACTTGGCAGTACATCTACGTATTAGTCATC  
 GCTATTACCATGGTATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGAC  
 TCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGAGTTGTTGGCACCAA  
 AATCAACGGGACTTCCAAAATGTCGTAACAACCTCCCTATGACGCAAATGGCGGT  
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCCTATCAGTGATAGAGATCTC  
 CCTATCAGTGTAGAGATCGTGACGGAGCTCGTTAGTGAACCGTCAGATGCCCTGGAGA  
 CGGCATCCACGCTGTTTGACCTCCATAGAACAGACACGGGACCGATCCAGCCTCCGGACT  
 CTAGAGGATCCCTACCGGTGATATCCTCGAGCCATCAACAAGTTGTACAAAAAGCTG  
 AACGAGAAACGTAAGATGATAAAATCAATATATTAAATTAGATTGATAGGAC  
 AGACTACATAACTGTAAAACACACATATCCAGTCACTATGGCGGCCGATTAGGCAC  
 CCCAGGCTTACACTTTATGCTCCGGCTGTATAATGTTGGATTGAGTAGGATCC  
 GGCAGAGTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCAC  
 CGTTGATATATCCAATGGCATCGTAAAGAACATTGAGGCATTTCAGTCAGTTGCTCA  
 ATGTACCTATAACCAGACCGTTAGCTGGATATTACGGCTTTAAAGACCGTAAAGAA  
 AAATAAGCACAAGTTTATCCGGCTTATTCAACATTCTGCCCGCTGATGAATGCTCA  
 TCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGTATGGGATAGTGTTCACCC  
 TTGTTACACCGTTTCCATGAGCAAACGTTTACAGTAAAGAACATTGAGGCTAGTCAGTTGCTCA  
 CGACGATTCCGGCAGTTCTACACATATATTGCAAGATGTGGCGTTACGGTAAAAA  
 CCTGGCCTATTCCCTAAAGGGTTATTGAGAATATGTTTCGTCAGCCAATCCCTG  
 GGTGAGTTTACCACTGGGAAATTATACGCAAGGGCACAAGGTGCTGATGCCGCTGGCGATTCA  
 TTTCACCATGGGAAATTATACGCAAGGGCACAAGGTGCTGATGCCGCTGGCGATTCA  
 GGTCATCATGCCGTCTGTGATGGCTTCCATGCGCAGAATGCTTAATGAATTACAACA  
 GTACTGGATGAGTGGCAGGGCGGGCGTAAAGATCTGATCCGGTTACTAAAAGCCAG  
 ATAACAGTATGCGTATTGCGCGCTGATTTGCGGTATAAGAATATATACTGATATGTA  
 TACCCGAAGTATGTCAAAAGAGGTGTGCTATGAGCAGCGTATTACAGTGACAGTTGAC  
 AGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGCTGGTAAGCA  
 CAACCATGCGAATGAAGCCGTCGTGCGCAACGCTGGAAAGCGGAAATCAGG  
 AAGGGATGGCTAGGTCGCCCCGTTATTGAAATGAACGGCTTTGCTGACGAGAACAA  
 GGGACTGGTAAATGCAGTTAACGTTACACCTATAAAAGAGAGAGGCCGTATCGTCTG  
 TTTGTGGATGTACAGAGTGATATTATTGACACGCCGGCGACGGATGGTGTACCCCTG  
 GCCAGTGCACGCTGCTGTCAGATAAAAGTCTCCCGTGAACCTTACCCGGTGTGCATATC  
 GGGGATGAAAGCTGGCGATGATGACCAACGATATGGCCAGTGTGCCGGTCTCCGTTATC  
 GGGGAAGAAGTGGCTGATCTGCCACCGCGAAAATGACATCAAAACGCCATTACCTG  
 ATGTTCTGGGAAATATAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTCGACCA  
 TAGTGACTGGATATGTTGTTACAGTATTATGATGTTCTGTTTATGCAAAATCTA  
 ATTTAATATATTGATATTATCATTACGTTCTCGTTCAGCTTCTGTACAAAGT  
 GGTTGATGGCGCCGCTCTAGAGGGCCAAGCTTACCGCGTGCATGCGACGTGTCAGCTC  
 TCTCCCTATAGTGAGTCGTATTATAAGCTAGGCACTGCCGTGTTTACAACGTCGTGA  
 CTGGAAAAGTCTAGCTGGATCTTGTGAGGAACCTACTTCTGTGGTGTGACATA  
 ATTGGACAAACTACCTACAGAGATTAAAGCTAAGGTAATATAAAATTAAAGTGT  
 ATAATGTGTTAAACTAGCTGCATATGCTTGTGCTTGAGAGTTGCTTACTGAGTATGA  
 TTTATGAAAATATTATAACACAGGGAGCTAGTGATTCTAATTGTTGTGATTAGATTCA  
 CAGTCCCAAGGCTCATTCAAGGCCCTCAGTCCTCACAGTCTGTCATGATCATAATCAG  
 CCATACCACATTGTAGAGGTTTACTGCTTAAAAACCTCCACACCTCCCCCTGAA  
 CCTGAAACATAAAATGAATGCAATTGTTGTTAACCTGTTATTGCACTTATAATGG  
 TTACAAATAAGCAATAGCATCACAATTCAAAATAAGCATTTTCACTGCTTC  
 TAGTGTGGTTGTCAAACACTCATCAATGTATCTTATCATGTCAGGATCGATCCTG  
 AATGAATGGCCAACGCCGGGGAGAGGCCGTTGCGTATTGGCTGGCGTAATAGCGAAG  
 AGGCCGCACCGATGCCCTCCCAACAGTGGCGCAGCTGAATGGCGAATGGGACGCC  
 CCTGTAGCGGCCTTAAGCGCGGGGTGTTACGCGCAGCGTACCGCTACAC  
 TTGCCAGCGCCCTAGGCCGCTCCTTGTGTTCTCCCTTCTGCCACGTTG  
 CCGGCTTCCCGTCAAGCTAAATCGGGGCTCCCTTAGGGTTCCGATTAGTGTCTT-

FIGURE 92B

211/240

TACGGCACCTCGACCCCCAAAAAACTTGATTAGGGTGTGATGGTTCACGTAGTGGGCCATCGC  
 CCTGATAGACGGTTTCGCCCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCT  
 TGTTCAAAACGGAAACAACACTCAACCTATCTCGGTCTATTCTTTGATTTATAAGGGA  
 TTTGCCGATTCGGCCTATTGGTAAAAAAATGAGCTGATTTAACAAATTTAACGCGA  
 ATTTAACAAAATATTAAACCTTACAATTTCGCCTGATGCGGTATTTCTCCTTACGCAT  
 CTGTGCGGTATTCACACCGCATACGCGGATCTGCGCAGCACCATGCCCTGAAATAACCT  
 CTGAAAGAGGAACCTGGTTAGGTACCTTCTGAGGCGGAAAGAACCGAGCTGTGGAATGTGT  
 GTCAGTTAGGGTGTGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC  
 ATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTA  
 TGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCCTCCCTAACTCCGCCATCC  
 CGCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCATGGCTGACTAATTTTTTA  
 TTTATGCAGAGGCCAGGCCGCTCGCCCTTGAGCTATTCCAGAAGTAGTGAAGGAGGCT  
 TTTTGAGGCCCTAGGCTTTGCAAAAGCTGATTCTCTGACACAACAGTCTCGAACT  
 TAAGGCTAGAGCCACCATGATTGAACAGATGGATTGCACGCAGGTTCTCCGCCGCTTG  
 GGTGGAGAGGCTATTGGCTATGACTGGGACAACAGACAATCGGCTGCTCTGATGCCGC  
 CGTGTCCGGCTGTCAGCGCAGGGGCCGGTCTTTTGTCAAGACCGACCTGTCCGG  
 TGCCCTGAATGAACTGCAGGACAGCAGGCCGCTATCGTGGCTGGCACGACGGCGT  
 TCCTTGCGCAGCTGTGCTCGACGTTGCACTGAAGCGGGAAAGGGACTGGCTGCTATTGGG  
 CGAAGTGCAGGGCAGGATCTCTGTCATCTCACCTGCTCCTGCCAGAAAGTATCCAT  
 CATGGCTGATGCAATGCGGGCTGCATACGCTTGTACCGGCTACCTGCCATTGACCA  
 CCAAGCGAAACATGCATCGAGCGAGCACGTACTCGGATGGAAGCCGTCTGTCGATCA  
 GGATGATCTGGACGAAGAGCATCAGGGGCTCGGCCAGCGAACTGTTGCCAGGCTCAA  
 GGCAGCATGCCGACGGCAGGATCTCGTGTGACCCATGGCGATGCCCTGCTTGCGAA  
 TATCATGGTGGAAAATGGCGCTTTCTGGATTGACTGTGGCCGGCTGGGTGTGGC  
 GGACCGCTATCAGGACATAGCGTTGGCTACCGTGTATTGCTGAAGAGCTGGCGGCGA  
 ATGGGCTGACCGCTCTCGTGTACGGTATCGCGCTCCGATTGCAAGCGCATCGC  
 CTTCTATGCCCTCTTGACGAGTTCTCTGACGGGACTCTGGGTTGAAATGACCGAC  
 CAAGCGACGCCAACCTGCCATCAGGATGGCGCAATAAAATATCTTATTTCATTACA  
 TCTGTGTGTGGTTTTTGCTGAATCGATAGCGATAAGGATCCGCTATGGTGCACTCT  
 CAGTACAATCTGCTCTGATGCCCATAGTTAACGCCAGCCCCGACACCGCCAACACCCGC  
 TGACGCGCCCTGACGGCTGTGCTGCCGATCGCCTTACAGACAAGCTGTGACCGT  
 CTCCGGAGCTGATGTGTCAGAGGTTTACCGTCATACCGAAACGCGCAGACGAAA  
 GGGCCTCGTGTACGCCCTATTAGGTTAATGTCATGATAATAATGGTTCTTAGAC  
 GTCAGGGTGGCAGTTCGGGAAATGTGCGCGAACCCCTATTGTTATTCTAAAT  
 ACATTCAAATATGATCCGCTCATGAGACAATAACCGTATAATGCTTCAATAATATTG  
 AAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCTATTCCCTTTGCGGC  
 ATTTCGCTTCTGTTGCTCACCAAGAACGCTGGTGAAGTAAAGATGCTGAAGA  
 TCAGTTGGGTGACGAGTGGTTACATCGAAGTGGATCTAACAGCGTAAGATCCTGA  
 GAGTTTCGCCCCGAAGAACGTTTCAATGATGAGCAGCTTTAACGTTCTGCTATGTGG  
 CGCGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTGCCGCATACACTATT  
 TCAGAATGACTGGTTGAGTACTCACCAAGTCAGAAAAGCATCTACGGATGGCATGAC  
 AGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAACACTCGGCCAACTTACT  
 TCTGACAACGATCGGAGGACGAAGGAGCTAACGCTTTTGCAACACATGGGGATCA  
 TGTAATCGCCTGATCGTGGAAACCGGAGCTGAATGAAGCCATACCAACGAGCG  
 TGACACCACGATGCCGTAGCAATGCAACACGTTGCCAAACTATTAACTGGCGAAGT  
 ACTTACTCTAGCTCCGCCAACAAATTAGACTGGATGGAGGGGATAAAAGTTGCAGG  
 ACCACTTCTCGCCTCGGCCCTCCGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGG  
 TGAGCGTGGGTCTCGCGGTATCATTGCAAGCACTGGGCCAGATGGTAAGCCCTCCG  
 CGTAGTTATCTACAGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCG  
 TGAGATAGGTGCCACTGATTAAGCATGGTAACTGTCAGACCAAGTTACTCATATA  
 ACTTTAGATTGATTAAAACCTCATTTAATTAAAAGGATCTAGGTGAAGATCCTTT  
 TGATAATCTCATGACCAAAATCCCTTAACGTCAGGTTTCTGTTCCACTGAGCGTCA  
 CGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTCTGCGCGTAATCTGCTGCT  
 GCAAACAAAAAAACCCACCGCTACCAAGCGGTGTTGTTGCCGGATCAAGAGCTAC  
 TCTTTTCCGAAGGTAACGGCTCAGCAGAGCGCAGATACCAAAACTGTCTTCTAGT  
 GTAGCCGTAGTTAGGCCACCTCAAGAACACTGTAGCACCGCTACATACTCGCTCT  
 GCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGG  
 CTCAAGACGATAAGTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTCGTGCAC-

FIGURE 92C

212/240

ACAGCCCAGCTGGAGCGAACGACCTACACCGAACCTGAGATACTACAGCGTGAGCATTG  
AGAAAGCGCCACGCCCTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT  
CGGAACAGGAGAGCGCACGAGGGAGCTTCAGGGGGAAACGCCTGGTATCTTATAGTCC  
TGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTGTGATGCTCGTCAGGGGGCG  
GAGCCTATGGAAAAACGCCAGCAACCGCCCTTTTACGGTTCTGGCCTTTGCTGGCC  
TTTGCTCACATGTTCTTCCTGCGTTATCCCCGATTCTGTGGATAACCGTATTACCGC  
CTTGAGTGAGCTGATACCGCTGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAG  
CGAGGAAGCGGAAGAGGCCAATACGCAAACGCCCTCCCCGCGCTGGCCGATTCA  
TTAATGCAGAGCTTGCATTGCGCGTTTCAATTATTGAAGCATTATCAGGGTTA  
TTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAGGGTTCC  
GCGCACATTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATT  
AACCTATAAAATAGGCGTAGTACGAGGCCCTTCACTCATTAG

FIGURE 92D

213/2ho

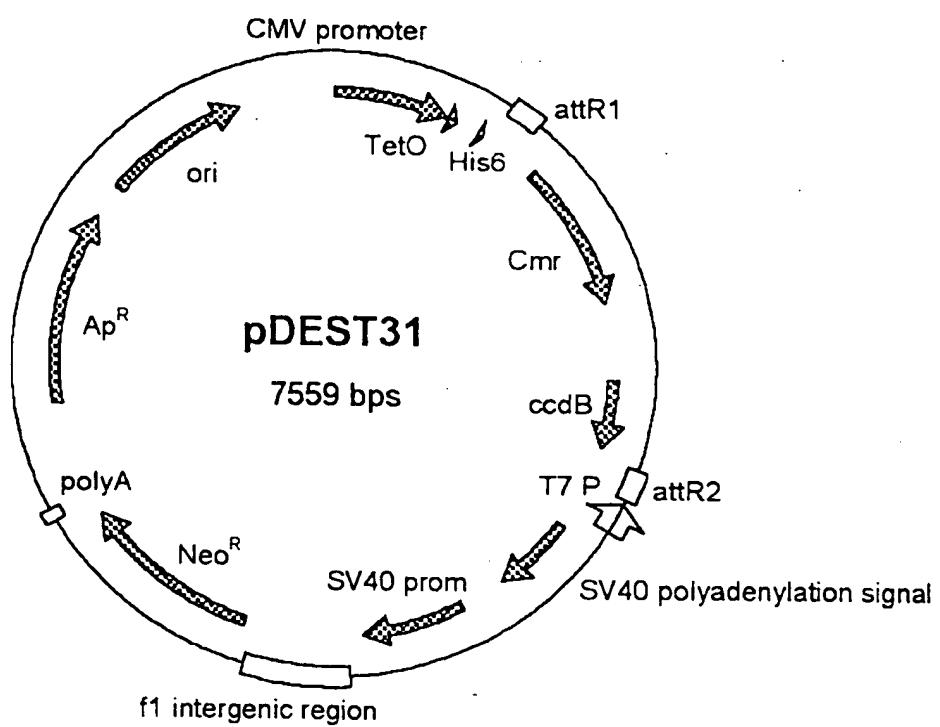


FIGURE 93A

214/240

pDEST31 7559 bp

ATGCATGTCGTTACATAACTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC  
 CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT  
 TGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTAT  
 CATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT  
 GCCCAGTACATGACCTATGGGACTTCCACTTGGCAGTACATCTACGTATTAGTCATC  
 GCTATTACCATGGTATGCCGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGAC  
 TCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGAGTTGTTGGCACCAA  
 AATCAACGGGACTTCCAAGTGTGTAACAACCTCCGCCATTGACGCAAATGGCGGT  
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCATACGTGATAGAGATCTC  
 CCTATCAGTGTAGAGATCGTGACGGAGCTCGTTAGTGAACCGTCAGATGCCCTGGAGA  
 CGGCCATCCACGCTGTTGACCTCCATAGAACGACACCGGAGCGATCCAGCCCTGGAGC  
 ATGGCGTACTACCATCACCATCACACCCGGTGTATCCTCGAGCCATTACAAGT  
 TTGTACAAAAAGCTGAACGAGAACGTAAAATGATATAAATATCAATATATAATTAG  
 ATTTGCATAAAAACAGACTACATAACTGTAAAACACACATATCCAGTCACTATGG  
 CGGCCGATTAGGCACCCAGGTTACACTTATGCTTCCGGCTCGTATAATGTGTGGA  
 TTTTGAGTTAGGATCCGGGAGATTTCAGGAGCTAAGGAAGCTAAAATGGGAAAAAA  
 TCACTGGATATAACCACCGTTGATATATCCATGGCATCGTAAAGAACATTGAGGCAT  
 TTCAGTCAGTGTCAATGTACCTATAACAGACCGTTAGCTGGATATTACGGCTTT  
 TAAAGACCGTAAAGAAAAAAGCACAAGTTTATCCGGCTTATTACACATTGCCCC  
 GCCTGATGAATGCTCATCCGGAAATCCGTATGGCAATGAAAGACGGTGGATAT  
 GGGATAGTGTACCCCTGTTACACCGTTTCCATGAGCAAACGTAAAGCTTTCATCGC  
 TCTGGAGTGAATACCACGACGATTCCGGCAGTTCTACACATATTCGCAAGATGTGG  
 CGTGTACGGTAAAACCTGGCTATTCCCTAAAGGGTTATTGAGAATATGTTTCG  
 TCTCAGCCAATCCCTGGGTGAGTTCACAGTTGATTAAACGTGGCAATATGGACA  
 ACTTCTCGCCCCGGTTTACCATGGGAAATATTACGCAAGGGCACAAGGTGCTGA  
 TGCCGCTGGCATTAGTTGATCATCGCTGTGATGGCTTCCATGTCGGCAGAATGC  
 TTAATGAAATTACAAACAGTACTCGCATGAGTGGCAGGGGGGGCGTAAACGGTGGATCCG  
 GCTTACTAAAAGCCAGATAACAGTATGCGTATTGCGCGCTGATTGGCGTATAAGAA  
 TATATACTGATATGTATAACCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTAT  
 TACAGTGACAGTTGACAGCGACAGCTATGTTGCTCAAGGCATATATGATGTCAATATC  
 TCCGGCTGGTAAGCACAAACATGCGAATGAAGCCGTCGTGCGTGCAGAACGCTGG  
 AAAGCCGAAAATCAGGAAGGGATGGCTGAGGTGCCCCGGTTATTGAAATGAACGGCTCT  
 TTTGCTGACGAGAACAGGGACTGGTGAATGCAAGTTAACGTTACACCTATAAAAGAGA  
 GAGCCGTTATCGTCTGTTGAGTACAGAGTGATATTATTGACACGCCGGCGACG  
 GATGGTGTACCCCTGGCCAGTGCACGTCTGCTGTCAGATAAGCTCCCGTGAACCTTA  
 CCCGGTGGTCATATCGGGATGAAAGCTGGCGCATGATGACCAACGATATGGCCAGTGT  
 GCGCGTCTCCGTTATCGGGAAAGAAGTGGCTGATCTCAGCCACCGGAAAATGACATCAA  
 AAACGCCATTAAACCTGATGTTCTGGGAATAAAATGTCAGGCTCCGTTACACAGCCA  
 GTCTGCAGGTGACCATAGTGACTGGATATGTTGTTACAGTATTATGAGTCTGTT  
 TTTTATGAAATCTAATTAAATGATATTATATCATTACGTTCTCGTTCACTGTT  
 CTTCTGTACAAAGTGGTGTGATGGCGCCGCTCTAGAGGGCCAAGCTTACGCGTGCAT  
 GCGACGTCTAGCTCTCCCTATAGTGAGTGTGTTATAAGCTAGGCACTGGCCGTCGT  
 TTTACAACGTCGTGACTGGAAAAGCTGCTAGCTGGATCTTGTGAAGGAACCTTACTT  
 CTGTTGTTGACATAATTGGACAAACTACCTACAGAGATTAAAGCTCTAAGGTAATAT  
 AAAATTGTTAAAGTGTATAATGTTAAACTAGCTGATATGCTTGCTGCTTGAGAGTTT  
 GCTTACTGAGTATGATTATGAAAATATTACACAGGAGCTAGTGATTCTAATTGTTG  
 TGTATTGTTAGATTACAGTCCCAAGGCTCATTCAAGGCCCTCAGTCCTCACAGTCGT  
 CATGATCATAATCAGCCATACCACATTGTAGAGGTTTACTTGCTTAAACCTCCC  
 ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTAACTTGTATT  
 TGCAGCTTATAATGGTTACAAATAAGCAATAGCATCACAAATTCAACAAATAAGCATT  
 TTTTCACTGCAATTCTAGTTGTTGCTCAAACACTCATCAATGTTATCTTATCATGTC  
 GATCGATCCTGCATTAGAATGGCCAACGCCGCGGGAGAGGGCGTTGCGTATTGGCT  
 GCGCTAATAGCGAAGAGGCCGACCGATGCCCTCCAAACAGTTGCGCAGCCTGAATG  
 CGCAATGGGACGCCCTGAGCGCGCATTAGCGCCGCGGGTGTGGTACGCGCA  
 CGCGTACACTTGCCAGGCCCTAGCGCCGCTCCTTCGCTTCCCTTCC  
 TTCTGCCACGTTGCCGGCTTCCCGTCAAGCTCTAAATCGGGGCTCCCTTAGGGT-

FIGURE 93B

215/240

TCCGATTAGTGTACGGCACCTCGACCCCCAAAAACTTGATTAGGGTGTGGTCAC  
 GTAGTGGGCCATCGCCCTGATAGACGGTTTCGCCCTTGACGTTGGAGTCCACGTTCT  
 TTAATAGTGGACTCTGTTCAAACGGAAACAACACTCAACCCTATCTGGTCTATTCTT  
 TTGATTATAAGGGATTTGCCGATTGCCCTATTGGTAAAAAATGAGCTGATTTAAC  
 AAATATTTAACCGCAATTAAACAAATAATTACGTTAACATTTCGCCGTGCGCAGGCTAT  
 TTTCTCCTACGCATCTGCGGTATTACACCCGATACCGGATCTGCGCAGCACCCT  
 GCCCTGAAATAACCTCTGAAAGAGGAACCTGGTAGGTTAGCTCTGAGGCGGAAAGAAC  
 AGCTGTGAATGTGTGTCAGTTAGGGTGTGAAAGTCCCAGGCTCCCAGCAGGAGAAC  
 GTATGCAAAGCATGCATCTAACATTAGTCAGCAACCCAGGTGTGAAAGTCCCCAGGCTCCC  
 CAGCAGGAGAAGTATGCAAAGCATGCATCTAACATTAGTCAGCAACCATAGTCCC  
 TAACCTCGCCCATCCGCCCTAACCTCGCCAGTCCGCCATTCTCGCCCATGGCT  
 GACTAAATTTTTATTATGCAAGAGGCCGAGGCCGCTGGCTCTGAGCTATTCCAGA  
 AGTAGTGAAGGAGGCTTTTGAGGCCCTAGGCTTTGCAAAAAGCTTGATTCTGTGACA  
 CAACAGTCTGAACCTAACGGCTAGAGCCACATGATTGACAAGATGGATTGACAGCAGG  
 TTCTCGGCCGCTTGGTGGAGAGGCTATTGCTATGACTGGGACAACAGACAATCGG  
 CTGCTCTGATGCCGCCGTGTCGGCTGTGAGCGCAGGGCGCCGGTTCTTGTCA  
 GACCGACCTGCCCCCTGAATGAACTGCAAGGAGGCCGAGGCTATCGTGGCT  
 GGCCACGACGGCGTTCGCAGCTGTGCTGACGTTGTCAGTGAAGCGGGAGGG  
 CTGGCTGCTATTGGCGAAGTGCCGGAGGATCTCTGTCATCTCACCTGCTCCTGC  
 CGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGCTGCATACGCTTGATCCGGCTAC  
 CTGCCATTGACCAACCGAAACATCGCATCGAGCAGCACGACTCGGATGGAAGC  
 CGGTCTGTCATCAGGATGATCTGGACGAAGAGCATCAGGGCTCGGCCAGCGAACT  
 GTTCGCCAGGCTCAAGGCGCATGCCGACGGCAGGATCTGTCGTGACCGATGGCGA  
 TGCCTGCTTGCAGGAAATATCATGGTGGAAAATGCCGCTTTCTGGATTGACTGTGG  
 CCGCTGGGTGGCGGACCGCTATCAGGACATAGCGTGGCTACCGTGATATTGCTGA  
 AGAGCTGGCGGAATGGCTGACCGCTTCTGCTTACGGTATGCCGCTCCCGA  
 TTCGCAAGCGCATGCCCTCATGCCCTTGTGACGAGTTCTGAGCGGGACTCTGGGG  
 TTCGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGATGCCGCAATAAAATATC  
 TTTATTTCATTACATCTGTGTTGGTTTGTGTAATCGATAAGCGATAAGGATCCG  
 CGTATGGTCACCTCAGTACAATCTGCTCTGATGCCGATAGTTAACGCCAGCCCCGACA  
 CCCGCCAACACCCGCTGACGCCCTGACGGGCTTGTGCTCCCGCATCCGCTTACAG  
 ACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTCAACGTCATACCGAA  
 ACGCGAGACGAAAGGGCTCGTGATACGCCATTGGTAAATGTGATGATAAT  
 AATGGTTCTTAGACGTCAAGTGGCACTTTGGGAAATGTGCGGGAACCCCTATTG  
 TTTATTTCATAACATTCAAATATGATCCGCTCATGAGACAATAACCCGATAAAT  
 GCTCAATAATATTGAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCTAT  
 TCCCTTTTGCGGCTTGCCTTCTGTTTGCTCACCGAGAACGCTGGTGAAGT  
 AAAAGATGCTGAAGATCAGTGGTGCACGAGTGGTTACATGAACTGGATCTCAACAG  
 CGTAAGATCCTGAGAGTTTGCCTCGAAGAACGTTCCAATGATGAGCACTTTAA  
 AGTTCTGCTATGCGCGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTG  
 CCGCATACACTATTCTCAGAATGACTGGTGTAGTACTCACAGTCACAGAAAAGCATCT  
 TACGGATGGCATGACAGTAAGAGAATTATGCACTGGTGTGCTGCCATAACCATGAGTGATAACAC  
 TGCGCCAACCTACTCTGACAACGATCGAGGACCGAAGGAGCTAACGCTTTTGCA  
 CAACATGGGGATCATGTAACCTGCCCTGATCGTGGGAACCGGAGCTGAATGAAGCCAT  
 ACCAACGACGACGCGTGACACCACGATGCCGTAGCAATGGCAACACGTTGCGCAA  
 ATTAACCTGGCAACTACTCTAGCTTCCCGCAACAATTAAATAGACTGGATGGAGGC  
 GGATAAAAGTTGAGGACCACTTCTGCGCTCGGCCCTCCGGCTGGTTATTGCTGA  
 TAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATTGACGCCGGCAAGAGCA  
 TAAGCCCTCCCGTATCGTAGTTATCACAGACGGGAGTCAGGCAACTATGGATGAACG  
 AAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGTAAGTGT  
 AGTTACTCATATATACTTTAGATTGATTAAACTTCATTAAATTTAAAGGATCTA  
 GGTGAAGATCCTTTTGATAATCTCATGACCAAATCCCTAACGTGAGTTCTGTTCA  
 CTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTCTGAGATCCTTTCTGCG  
 CGTAATCTGCTGCTTGCACCAACAAAAACCCACCGCTACAGCGGTGGTTGTTGCC  
 TCAAGAGCTACCAACTCTTCCGAAGGTAACGGCTTCAGCAGAGCGCAGATACCA  
 TACTGCTCTTAGTGTAGCCGTAGTTAGGCCACCAACTCAAGAACTCTGAGCACC  
 TACATACCTCGCTCTGCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATA  
 TCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTGGGCTGAAC-

Figure 93c

GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAACCT  
ACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCGAAGGGAGAAAGCGGGACAGGTATCC  
GGTAAGCGGCAGGGTCGGAACAGGGAGAGCGCACAGGGAGCTTCCAGGGGAAACGCCTG  
GTATCTTATAGTCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTGTGATG  
CTCGTCAGGGGGCGGAGCTATGGAAAAACGCAGCAACCGCCCTTTTACGGTTCCCT  
GGCCTTTGCTGGCCTTTGCTCACATGTTCTTCCTGCGTTATCCCTGATTCTGTGGA  
TAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCG  
CAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCCCCGC  
GCGTTGGCCGATTCTTAATGCAGAGCTTGCATTGCGCGTTTCAATATTATTGAAG  
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGTATTTAGAAAAATAA  
ACAAATAGGGTTCCGCGCACATTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCAT  
TATTATCATGACATTAACCTATAAAATAGGCGTAGTACCGAGGCCCTTCACTCATTAG

FIGURE 93D

217/240

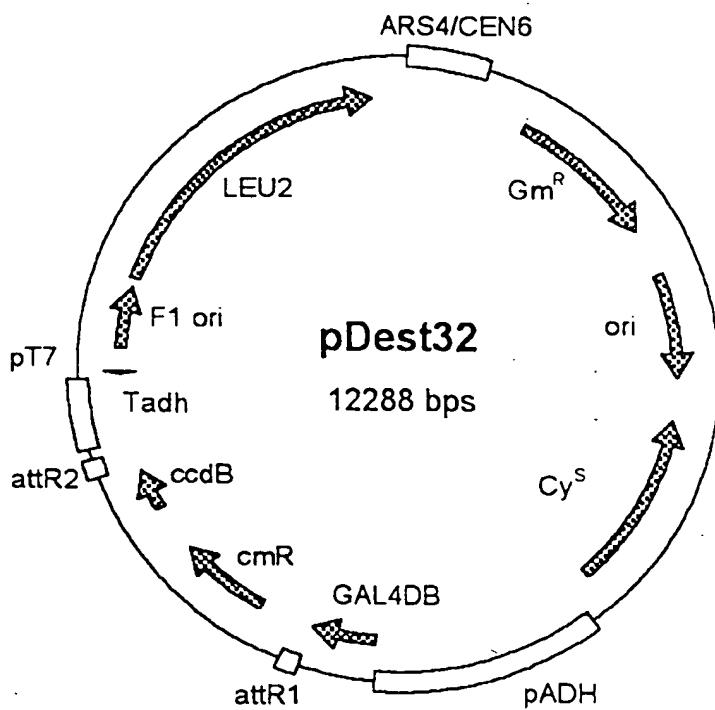


FIGURE 94A

pDEST32 12288 bp

GACGAAAGGGCCTCGT GATA CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTT  
 CTTAGGACGGATCGCTGCCTGTAACCTACACCGCCTGTATCTTTAATGATGGAATA  
 ATTTGGAATTACTCTGTTATTTATTTATGTTGTATTTGGATTTAGAAAGT  
 AAATAAAGAAGGTTAGAAGAGTTACGGAATGAAGAAAAAAAATAAACAAAGGTTAAAAA  
 ATTTCAACAAAAGCGTACTTACATATATTTATTAGACAAGAAAAGCAGATTAAATA  
 GATATACATTGATTAACGATAAGTAAATGTAACAGGATTTCGTGTGGTCT  
 TCTACACAGACAAGATGAAACAATTCCGCATTAATACCTGAGAGCAGGAAGAGCAAGATA  
 AAAGGTTAGTATTGTCGGCATCCCCTAGAGCTTTACATCTCGAAAACAAAAACT  
 ATTTTTCTTAATTCTTTTACTTCTATTTAATTATATTTATTTATTTATTTAAAAA  
 ATTTAAATTATAATTATTTATAGCACGTGATGAAAAGGACCCAGGTGGCACTTTGG  
 GGAAATGTGCGGGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCCG  
 CTCATGAGACAATAACCCTGATAAAATGCTTCAATAATCTGAGTGCAGCAGGGCCGTGTC  
 TCAAAATCTCTGATGTTACATTGACAAGATAAAAATATCATCATGAACAATAAAACT  
 GTCTGCTTACATAAACAGTAATAACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTC  
 TTGCTGGAGGCCGATTAAATTCAAACATGGATGCTGATTATATGGGTATAATGGC  
 TCGGTAGCCAACCACTAGAACTATAGCTAGAGTCTGGCGAACAAACGATGCTCGCCTT  
 CCAGAAAACCGAGGATGCGAACCACTTCATCCGGGTCAAGCACCACGGCAAGCGCCGCG  
 ACGGCCGAGGTCTCCGATCTCTGAAGCCAGGGCAGATCCGTGACAGCACCTGCG  
 AGAAGAACAGCAAGGCCAATGCTGACGATGCGTGAGACCGAAACCTTGCCTCG  
 TCGCCAGCCAGGACAGAAATGCCCTGACTTCGCTGCTGCCAAGGGTGCCTGACGCA  
 CACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATAAGCCTGTTGGTCTGTAAC  
 TGTAATGCAAGTAGCGTATGCGCTACGCAACTGGTCCAGAACCTTGACCGAACGCG  
 GTGGTAACGGCGAGTGGCGTTTCATGGCTGTTATGACTGTTTTGTACAGTCTA  
 TGCCTGGCATCCAAGCAGCAAGCGCTTACGCCGTGGTCGATGTTGATGTTATGGA  
 GCAGCAACGATGTTACGCGAGCAACGATGTTACGAGCAGGGCAGTCGCCCTAAAACA  
 AAGTTAGGTGGCTCAAGTATGGCATCATTGCCACATGTAGGCTGGCCCTGACCAAGTC  
 AAATCCATGGGGCTGCTTGTATCTTGTGAGTCTGGCTGAGTTGGAGACGTAGCCACCTAC  
 TCCCACATGCCGGACTCCGATTACCTGGAACTTGCTCCGTAGTAAGACATTCA  
 GCGCTGCTGCCCTCGACCAAGAACGGTTGTTGGCGCTCTCGCGCTTACGTTCTGCC  
 AGGTTGAGCAGCCGCTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCAGCAC  
 CGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCTCAAGCATGAGGCCAACGCG  
 GGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGAGTGGCTCTCTAT  
 ACAAAAGTTGGCATACTGGGAGAACGGTGTGACTTTGATATGACCCAAAGTACCGCCACC  
 TAACAATTGTTCAAGCCGAGATCGGCTCCGGCTAATAGGTTGATTGATGTTGGAC  
 GAGTCGGAAATCGCAGACCGATACCAAGGATCTGCCATCCTATGGAACGCTCGGTGAGT  
 TTTCTCTTCAATTACAGAAACGGTTTTCAAAATATGGTATTGATAATCTGATATGA  
 ATAAATTGCAAGTTTCAATTGATGCTCGATGAGTTTTCTAATCAGAATTGGTAATTGGT  
 TGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGNCATGACCAAAATCCCT  
 AACGTGAGTTTCTGTTCACTGAGCGTCAGACCCGTAGAAAGATCAAAGGATCTCTT  
 GAGATCCTTTCTGCGCTAATCTGCTGCTTGCACAAACAAAAACCGCTACCA  
 CGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTCTGCAAGGTAACGGCTTCA  
 GCAGAGCGCAGATAACCAACTGCTCTAGTGAGCCGTTAGGCAACCAACTTCA  
 AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCGAGTGGCTG  
 CCAGTGGCGATAAGTCGTTACCGGGTGGACTCAAGACGATAGTTACGGATAAGG  
 CGCAGCGGTGGCTGAACGGGGGTTCTGACACAGCCCAGCTGGAGCGAACGACCT  
 ACACCGAAGTACGACCTACAGCGTGAGCATTGAGAACGGCCACGCTCCGAAGGG  
 GAAAGGCGGACAGGTATCCGTAAGCGCAGGGTGGAAACAGGGAGCGCACGAGGGAGC  
 TTCCAGGGGGAAACGCTGGTATCTTATAGTCCTGCTGGTTGCTGCCACCTCTGACTTG  
 AGCGTCGATTTGTGATGCTCGTCAGGGGGCGAGGCTATGGAAAAACGCCAGCAACG  
 CGGCCTTCTACGGTCTGGCCTTTGCTGGCTTTGCTCACATGTTCTCTGCGT  
 TATCCCTGATTCTGTGATAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTGCC  
 GCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATAC  
 GCAAACCGCCTCTCCCGCGCTGGCCGATTCAATTAAATGCAAGCTGGCACGACAGGTTTC  
 CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAAATGAGTTACCTCACTCATTAGG  
 CACCCCAAGGCTTACACTTTATGCTTCCGGCTCTATGTTGTGGAATTGAGCGGAT  
 AACAAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGGAATTAAACCTC-

FIGURE 9B

219/26

ACTAAAGGGAACAAAAGCTGGTACCGATCCCGAGCTTGCAAAATTAAAGCCTCGAGCGT  
 CCCAAAACCTCTCAAGCAAGGTTTCAGTATAATGTTACATGCGTACACGCGTCTGTAC  
 AGAAAAAAAAGAAAATTGAAATATAAAACGTTCTTAATAACTAACATAACTATAAAA  
 AAATAAAATAGGGACCTAGACTTCAGGGTGTCTAACCTCCTTCGGTTAGAGCGGAT  
 GTGGGGGGAGGGCGTGAATGTAAGCGTACATAACTAATACATGATATCGACAAAGGAA  
 AAGGGGCTGTTACTCACAGGCTTCAAGTAGGTAAATTAGTCGTTCTGTCTTT  
 TCCTCTCAACCCACCAAGGCCATCTGGTACTTTTTTTTTTTTTTTTTTTTTTT  
 TTT  
 TTTTTTTCTAGAAAATAACAGAAAGTAGATGTTGAATTAGATTAAACTGAAGATATAT  
 AATTATGGAAAATACATAGAGCTTTGATGCGCTAACGATCAATTCAACAAAC  
 ACCACCAGCAGCTTGATTCTTCAGCCAACGGAGAATCTAGCTTGACGAT  
 AACTGGAACATTGGAATTCTACCCCTAACCAAGATCTAACGTAACCGCTGCCAAGT  
 GTCAATAACTGGAGCAGTTCTCTAGAACGAGATTCAAGTATTGGTCTCTGTCTTC  
 TGGGATCAATGTCACAATTGTCAGTTCAAGACTGGCTTCCAGAAATGAGCTTGTG  
 CTTGTGGAAGTATCTCATACCAACCTTACCGAAATAACCTGGATGGTATTATCCATGTT  
 AATTCTGTGGTGTGTTGACCAACGGGCATACCTCTAACCAACGGGGTGTCTGTGCTT  
 ACCGATAACGACCTTACCGCTGAGACGTGACCTCTGTGCTTCTAGTCTTAGTGAATCT  
 GGAAGGCATTCTGATTAGTGGATGATTGTTCTGGGATTTAATGCAAAATCACTTAAG  
 AAGGAAAATCAACGGAGAAAGCAAACGCCATCTAAATATAACGGGATACAGATGAAAGGG  
 TTTGAAACCTATCTGGAAAATAGCATTAACAAACGAAAAACTGCAGGGAAATTGTTGC  
 GTCTCTGCGGCTATTACCGGCCAGAGGAAAATAGGAAAATAACAGGGCATTAGAAAA  
 ATAATTGATTGATTTGGTAAATGTTGGGCTCTGGTGTACAGATGTTACATTGGTACAGTA  
 CTCTGTTTGTGTTGATGAAATCTCAAAATGGTGTAGCACATGGAAGAG  
 TCACCGATGCTAAGTTATCTATGTAAGCTACGTGGCGTACTTTGATGAAGCCGAC  
 AAGAGATAACGGATTGGCAACTGCAAATAGAAATCTGGGATCCCCCTCGAGATCCGGGA  
 TCGAAGGAAATGTTGGTAAATGAAATAGGAAATCAAGGACATGAAGGCAAAAGACAAATA  
 TAAGGGTCAACGAAAATAAGTAAAAGTGGGATGTTGATATGATGTTGGCTTGCAGCG  
 CCGAAAAAACGAGTTACCGCAATTGACAATCATGCTGACTCTGGCGGACCCGCGCTC  
 TTGCGGCGGCGATAACGCTGGCGTGGCTGAGGCTGTGCCGGAGTTTGCCTG  
 CATTTCAGGTTACCTCGCTAACGGGGAGATTGGAGAACAAATAAGAATGCCGG  
 TTGGGTTGCGATGATGACGACCACGACAACGGTGTGATTAAAGTTGCGGAAAGAA  
 CCTGAGTGCATTGCAACATGAGTAACTAGAAGAACGACTTGCGAGACGCGA  
 GTTGGCGGTGGTGCAGAACAAATAGAGCACCAGACCTGAAAGTGGAGACGCGCATAACC  
 GCTAGAGTACTTGAAGAGGAAACAGCAATAGGTTGCTACCAGTATAATAGACAGGTA  
 CATAACAACACTGAAATGGTGTCTGGTACGCTTCAATTCAATTGGGTGTGAC  
 TTTATTATGTTACAATATGGAAGGGAACTTACACTTCTCTATGCAACATATAATTAA  
 AAGTCCAATGCTAGTAGAGAACGGGGTAACACCCCTCCGCGCTCTTCCGATT  
 CTAACCGTGGAAATTTCGGATACCTTGTGTTCCGGGTGTACAATATGGACTTC  
 CTCTTTCTGGCAACCAACCCATACATCGGATTCTATAATACCTCGTGGTCTCCC  
 TAACATGTAGGTGGCGGAGGGGAGATACAAATAGAACAGATACCAAGACATAATG  
 GGCTAAACAAGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAAT  
 ACTGTAGCCCTAGACTTGTAGCCATCATATCGAAGTTCACTACCCCTTTCCATT  
 TGCCATCTATTGAAAGTAAATAGGCGCATGCAACTTCTTTCTTTCTTCTC  
 TCTCCCCGTTGTTGTCACCATACTCGCAATGACAAAAAAATGATGGAAGACACTAA  
 AGGAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTCCAGAGCTGATGAGG  
 GGTATCTCGAACACACGAAACTTTCTCTCATTGATGACACTACTCTCTAATG  
 AGCAACGGTATACGGCCTCCTCAGTTACTGAAATTGAAATAAAAAGTTGCCGC  
 TTTGCTATCAAGTATAAAAGACCTGCAATTATAATCTTTGTTCCCTCGTCAATTGTT  
 TCGTCTCTTCTCGTGTGTTCTGACAATATTCAGCTACCAAGCATAAC  
 AATCAACTCCAAGCTTGAAGCAAGCCTCTGAAAGATGAAAGCTACTGCTTCTATCGAAC  
 AAGCATGCGATATTGCCGACTAAAAAGCTCAAGTGTCCAAAGAAAAACCGAAGTGC  
 CCAAGTGTCTGAAGAACAAACTGGAGTGTGCGTACTCTCCAAAACCAAAAGGTCTCCGC  
 TGACTAGGGCACATCTGACAGAAGTGGAAATCAAGGCTAGAAAGACTGGAACAGCTATT  
 TACTGATTCTCGAGAACCTGACATGATTTGAAATGATTCTTACAGGATA  
 TAAAAGCATTGTTAACAGGATTATTGTTACAAGATAATGTAATAAGATGCCGTACAG  
 ATAGATTGGCTCAGTGGAGACTGATATGCCCTAACATTGAGACAGCATAGAATAAGTG  
 CGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGTGA  
 GGTGAAATCAAACAAGTTGTACAAAAAGCTGAACGAGAACGTAAGGATATAATA-

Figure 94c

270/240

TCAATATAATTAAATTAGATTTCGATAAAAAACAGACTACATAATACTGTAAAACACAAC  
 ATATCCAGTCACTATGGCGGCCGCTAAGTTGGCAGCATCACCGACGCACTTTGCACG  
 ATAAATACCTGTGACGGAAGATCACTTCGAGAATAAATAAATCCTGGTGTCCCTGTTGA  
 TACCGGGAGGCCCTGGGCCACTTTGGCAAATGAGACGTTGATCGGCACGTAAGAGG  
 TTCCAACTTTACACATAATGAAATAAGATCACTACCGGGCGTATTTTGAGTTATCGAG  
 ATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGA  
 TATATCCAATGGCATCGTAAAGAACATTGGAGGCATTTCAGTTGCTCAATGTAC  
 CTATAACCAGACCGTTCAGCTGGATATTACGGCTTTAAAGACCGTAAGAAAAATAA  
 GCACAAGTTTATCCGGCCTTATTACACATTCTGCCCGCTGATGAATGCTCATCCGGA  
 ATTCCGTATGGCAATGAAAGACGGTGAGCTGGTATATGGGATAGTGTTCACCTTGT  
 CACCGTTTCCATGAGCAAACCTGAAACGTTTCATCGCTCTGGAGTGAATACACGACGA  
 TTTCCGGCAGTTCTACACATATATTGCAAGATGTGGCGTGTACGGTGAAAACCTGGC  
 CTATTTCCCTAAAGGGTTATTGAGAATATGTTTCTGCTCAGCCAATCCCTGGGTGAG  
 TTTCACCACTTTGATTTAAACGTGGCAATATGGACAACCTTCTGCCCGGCGTTTCA  
 CATGGGCAAATATTATACGCAAGGCAGACAGGTGCTGATGCCGCTGGCATTAGGTCA  
 TCATGCCGCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTG  
 CGATGAGTGGCAGGGCGGGCGTAATCTAGAGGATCCGGCTACTAAAGCCAGATAACA  
 GTATGCCGTATTGCGCGCTGATTTGCGGTATAAGAATATATACTGATATGTATACCCG  
 AAGTATGTCAAAAAGAGGTGCTGATGAAGCAGCGTATTACAGTGACAGTTGACAGCGAC  
 AGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGCTGGTAAGCACAACCA  
 TGCAGAATGAAGCCCCTGCTGCGTGCCTGGAAACGCTGGAAAGCGGAAAATCAGGAAGGG  
 TGGCTGAGGTGCGCCGGTTATTGAAATGAACGGCTCTTGCTGACGAGAACAGGGACT  
 GGTGAAATGCAGTTAAGGTTACACCTATAAAAGAGAGAGGCGTTATCGTCTGTTGT  
 GATGTAAGAGTGTATTATTGACACGCCGGCGACGGATGGTGTACCCCCCTGGCAGT  
 GCACGCTGCTGTCAGATAAGTCTCCGTGAACTTACCCGGTGGTGTACATCGGGGAT  
 GAAAGCTGGCGCATGATGACCACCGATATGCCAGTGTGCCGGTCTCGTTATCGGGGAA  
 GAAAGTGGCTGATCTCAGGCCACCGGAAAATGACATCAAAACGCCATTAAACCTGATGTT  
 TGGGAATATAAATGTCAGGCTCCCTATACACAGCCAGTCTGCAGGTCGACCATAGTGA  
 CTGGATATGTTGTTACAGTATTATGTTAGTCTGTTTATGCAAAATCTAATTAA  
 TATATTGATATTATATCATTTACGTTCTCGTTCAGCTTCTGTACAAAGTGGTTG  
 ATGGCCGCTAAGTAAGTAAGACGTCGAGCTCTAAGTAAGTAACGCCGCCACCGCGGTGG  
 AGCTTGGACTTCTCGCCAGAGGTTGGTCAAGTCTCAAGGTTGCTGGCTTGTC  
 TACCTTGCCAGAAATTACGAAAAGATGGAAAAGGGTCAAATGTTGGTAGATACGTTGT  
 TGACACTCTAAATAAGCGAATTCTTATGATTATGATTATTATTAAATAAGTT  
 AAAAATAAGTGTATACAAATTAAAGTGAATCTTAGGTTTAAACGAAAATTCTT  
 GTTCTTGAGTAACTCTTCTGTTAGGTCAGGTTGCTTCAGGTATAGCATGAGGTCGC  
 TCTTATTGACCACACCTCTACCGCATGCCAGCAATGCCATCGCTCCCAATT  
 CACCCATTGAGATATGCTAACTCCAGCAATGAGTTGATGAATCTGGTGTATT  
 TGTCTCAGAGGACAATACCTGTTGTAATGTTCTTCCACACGGATCCCAATTGCCC  
 TAGTGAAGTCGTTACAACTCACTGCCGTCGTTTACAACGTCGACTGGAAAACCC  
 TGGCGTTACCAACTTAATGCCCTGCGACATCCCCCTTCGCCAGCTGGCGTAATAG  
 CGAAGAGGCCGACCGATGCCCTTCCAAACAGTTGCGCAGCCTGAATGGCGAATGGAC  
 GCGCCCTGTAGCGCGCATTAAGCGCGGGGTGTGGTGTACGCGCAGCGTACCGCT  
 ACACCTGCCAGGCCCTAGGCCGCTCTTCGTTCTCCCTTCGCCACG  
 TTCGCCGGCTTCCCGTCAAGCTCTAAATGGGGGCTCCCTTAGGGTTCCGATTAGT  
 GCTTACCGCACCTGACCCAAAAACTGATTAGGGTGTGGTCACTGAGTGGCCA  
 TCGCCCTGATAGACGGTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGG  
 CTCTTGTCCAAACTGGAACAAACACTCAACCCCTATCTCGGTCTATTCTTTGATTAA  
 GGGATTGCGCTATTGGTAAAAATGAGCTGATTAAACAAAATTAAAC  
 GCGAATTAAACAAAATTAAACGTTACAAATTCTGATGCCGTTAGTTCTACGC  
 ATCTGTGCGGTATTACACCGCATATCGACGGCTGAGGAGAAACTCTAGTATATCCAC  
 ATACCTAATTATTGCTTATTAAAAATGGAATCGGAACAAATTACATCAAACATCCAC  
 TCTCTTCAAAATCAATTGCTCTGACTCTCTGTTCATGTTGTTCAAAACGTTATATT  
 TATAGGATAATTATACTCTATTCTCAACAAAGTAATTGGTTGTTGGCCGAGCGGTCTAA  
 GGCGCCGATTCAAGAAATCTTGACCGCAGTTACTGTGGGAAACTCAGGTATCGTA  
 AGATGCAAGAGTTCGAATCTCTAGCAACCATTATTTTCTCAACATAACGAGAAC  
 CACAGGGCGCTATGCCACAGAATCAAATTGATGACTGGAAATTGGTTAATTTCAG  
 AGGTGCCGCTGACCGCATACCTTTCAACTGAAAATTGGAGAAAAGGAAAGGTGAG-

FIGURE 94D

221/260

AGGCCGGAACCGGCCTTTCATATAGAATAGAGAACGGTTCATGACTAAATGCTTGCATCA  
CAATACTTGAAGTTGACAATATTATTAAAGGACCTATTGTTTTCAATAGGTGGTTAG  
CAATCGTCTTAACCTTCTAACCTTCTACCTTACATTTCAGCAATATATATATATT  
TCAAGGATATACCAATTCTAATGTCTGCCCTATGTCTGCCCTAAGAACATCGTCGTTT  
GCCAGGTGACCACGGTCAAGAAATCACAGCCGAAGCATTAAAGGTTCTAAAGCTAT  
TTCTGATGTTGCTCCAATGTCAGGTCGATTCGAAAGGCTTAAAGGTTCTAAAGCTAT  
TATCGATGCTACAGGTGTCCTCCAGATGAGGCGCTGGAAAGCCTCCAAGAACGGTTGA  
TGCGTTTGTAGGTGCTGGGTGGCTCTAAATGGGGTACCGGTAGTGTAGACCTGA  
ACAAGGTTTACTAAAAATCCGAAAGAACCTCAATTGTACGCCAACCTAACGACCATGTAA  
CTTGCATCCGACTCTTTAGACTTATCTCAATCAAGCCACAAATTGCTAAAGGTAC  
TGACTTCGTTGTCAGAGAATTAGTGGGAGGTATTTACTTGGTAAGAGAACAGGAAGA  
CGATGGTGTGGTGTGCTGGGATAGTGAACAATACACCGTTCAGAACGTGCAAAGAAAT  
ACAAGAAATGCCGCTTCATGCCCTACACATGAGCCACCATTGCCTATTGGCCTT  
GGATAAAAGCTAATGTTTGGCCTTCAAGATTATGGAGAAAAGTGTGGAGGAAACCAT  
CAAGAACGAATTCCCTACATTGAAGGTTCAACATCAATTGATTGATTCTGCCGCCATGAT  
CCTAGTTAAGAACCCAACCCACCTAAATGGTATTATAATCACCAGCAACATGTTGGTGA  
TATCATCTCCGATGAAGCCTCCGTTATCCCAGGTTCTGGTTGTGCCATCTCGTC  
CTTGGCCTTTGCCAGACAAGAACACCGCATTTGGTTGTACGAACCATGCAACGGTTC  
TGCTCCAGATTGCCAAAGAATAAGGTTGACCCCTATGCCACTATCTGTCTGCTGCAAT  
GATGTTGAAATTGTCATTGAACTTGCCTGAAGAACGGTAAGGCCATTGAAGATGCAAGTTAA  
AAAGGTTTGGATGCGAGGTATCAGAACTGGTGTAGTTAGGTGGTCCAACAGTACCAACCGA  
AGTCGGTGTGTCGCCAGAACAGTTAAGAAAATCCTGCTAAAAAGATTCTCTTTT  
TTTATGATATTGTCATAAAACTTTATAATGAAATTCTATAATAGAAAACGACACGAAATT  
ACAAAATGGAATATGTCATAGGTAGACGAAACTATATAACGCAATCTACATACATTAT  
CAAGAAGGAGAAAAGGAGGATAGTAAAGGAATACAGGTAAGCAAATTGATACTAATGGC  
TCAACGGTATAAGGAAAAAGAATTGCACTTAACTTAAATTGACAAGGAGGGAGGGCAC  
CACACAAAAAGTTAGGTGTAAACAGAAAATCATGAAACTACGATTCTAAATTGATATTGG  
AGGATTCTCTAAAAAAAAAAATACAACAAATAAAAACACTCAATGACCTGACCAT  
TTGATGGAGTTAAGTCATAACCTTCTGAACCATTCCATAATGGTGAAGTTCCCTC  
AAGAATTCTACTCTGTCAGAAACGGCTTACGACGTAGTCGATATGGTCACCTCTCAGTA  
CAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCAACACCGCTGACG  
CGCCCTGACGGGCTTGTCTGCTCCGGCATCCGTTACAGACAAGCTGTGACCGTCTCCG  
GGAGCTGCATGTGTCAGAGGTTTCACCGTCATACCGAAACGCGCGA

FIGURE 94E

222/240

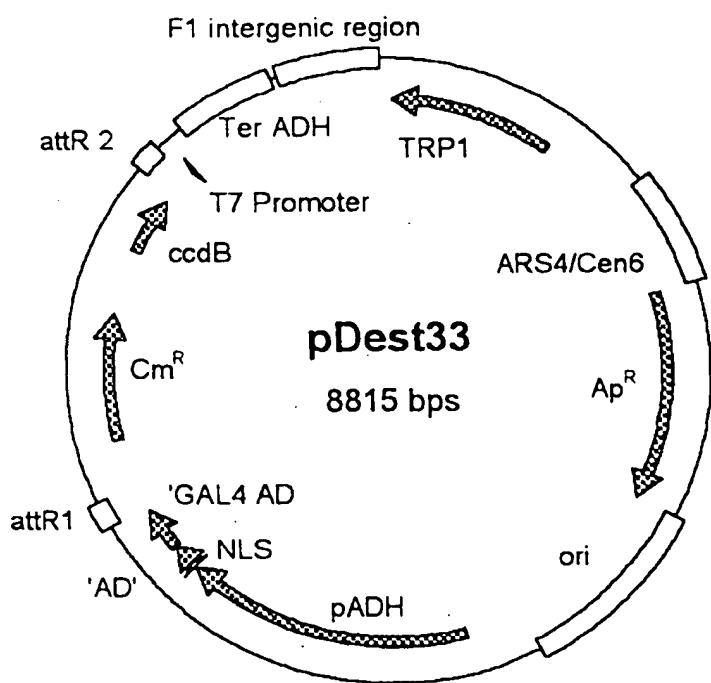


FIGURE 95A

223/240

pDEST33 8815 bp

GCCTTACGCATCTGTGCGGTATTCACACCGCAGGCAAGTCACAAACAATACTTAAATA  
 AATACTACTCAGTAATAACCTATTCTTAGCATTTCGACGAAATTGCTATTGTTAG  
 AGTCTTTACACCATTGCTCCACACCTCCGCTTACATCAACACCAATAACGCCATT  
 ATCTAACCGCATCACCAACATTCTGGCGTCAGTCCACCAAGCTAACATAAAATGTAAGC  
 TTTCGGGGCTCTCTGCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCAC  
 CTGTCCCACCTGCTCTGAATCAAACAGGGATAAACGAATGAGGTTCTGTGAAGCTG  
 CACTGAGTAGTATGTTGCAGTCCTTGGAAATACGAGTCCTTAATACTGGCAAACCGA  
 GGAACCTTGGTATTCTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGT  
 AATCATTGACCAGAGCCAAACATCCTCCTAGGTTGATTACGAAACACGCCAACCAAGT  
 ATTCGGAGTGCCCTGAACATTTCATTTATGCTTTACAAGACTTGAATTTCTTGCA  
 TAACCGGGTCAATTGTTCTTTCTATTGGGACACATATAATACCCAGCAAGTCAGCAT  
 CGGAATCTAGAGCACATTCTGGCCCTCTGTGCTCTGCAAGCCGAAACTTCACCAATG  
 GACCAGAACTACCTGTGAATTAAACAGACATACTCCAAGCTGCCCTTGCTGCTTAA  
 TCACGTATACTCACGTGCTCAATAGTCACCAATGCCCTCCCTTGCCCTCTCCTTT  
 TTTTCGACCGAATTAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGT  
 ACGTAAGGTGACAAGCTATTTCATAAAAGAATATCTTCACTACTGCCATCTGGCGTC  
 ATAAC TGCAAAGTACACATATATTACGATGCTGCTATTAAATGCTCCTATATTATATA  
 TATAGTAATGTCGTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAA  
 GCCAGCCCCGACACCCGCCAACACCCGCTGACCGGCCCTGACGGGCTTGCTGCTCCGG  
 CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTAC  
 CGTCATCACCAGAAACCGCGAGACGAAAGGGCTCTGTGATACGCCATTTCAGGTTA  
 ATGTGATGATAATAATGGTTCTTAGGACGGATCGCTGCTGTAACACCGCCCTC  
 GTATTTAATGATGGAATAATTGGGAAATTACTCTGTTATTATTTATGTT  
 TGTATTGGATTAGAAAGTAAATAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAA  
 AAATAAAACAAAGTTAAAAAATTCAACAAAAGCGTACTTACATATATATTAG  
 ACAAGAAAAGCAGATTAAATAGATATACATTGATTAACGATAAGTAAAATGAAAATCA  
 CAGGATTTCGTTGTTCTACACAGACAAGATGAAACAATTGGCATTAAACCT  
 GAGAGCAGGAAGAGCAAGATAAAAGGTAGTATTGTTGGCATTCCCTAGAGTCTTTA  
 CATCTCGGAAACAAAACATTTCCTTAATTCTTTTACTTCTATTAA  
 TTTATATATTATTTAAATTATAATTATTTATAGCACGTGATGAAAG  
 GACCCAGGTGGCACTTTCGGGAAATGTGCGCGAACCCCTATTGTTATTCTAA  
 ATACATTCAAATATGATCCGCTCATGAGACAATAACCTGATAATGCTTCAATAATAT  
 TGAAAAGGAAGAGTATGAGTATTCAACATTCCGCTGCGCCCTTATTCCCTTTGCG  
 GCATTTCGCTTCTGTTTGCTCACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAA  
 GATCAGTTGGGTGACGAGTGGTTACATCGAATGGATCTCAACAGCGTAAGATCCTT  
 GAGAGTTTCGCCCCGAAGAACGTTTCAATGATGAGCACTTTAAAGTTCTGCTATGT  
 GGCGCGTATTATCCCGTATTGACGCCGGCAAGAGCAACTCGGTCGCCGATACACTAT  
 TCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTACGGATGGCATG  
 ACAGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAACACTGCGGCCACTTA  
 CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACGCTTTTCAACACATGGGGAT  
 CATGTAACTCGCCTGATCGTGGGAAACCGGAGCTGAATGAAGCCATACCAACGACGAG  
 CGTGACACCACGATGCCGTAGCAATGGCAACACGTTGCGCAAACACTATTAACTGGCGAA  
 CTACTTACTCTAGCTCCCGCAACAATTAAATAGACTGGATGGAGGCGGATAAAGTTGCA  
 GGACCACTCTCGCTCGGCCCTCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCC  
 GGTGAGCGTGGGTCTCGGGTATCATTGCACTGGGCCAGATGGTAAGCCCTCCCGT  
 ATCGTAGTTATCACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATC  
 GCTGAGATAGGTGCTCACTGATTAAGCATTGTAACTGTCAGACCAAGTTACTCATAT  
 ATACTTAGATTGATTAAAATTCAATTAAATTAAAAGGATCTAGGTGAAGATCCTT  
 TTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTCTGTTCACTGAGCGTCAGAC  
 CCCGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTCTGCGCGTAATCTGCTGC  
 TTGCAAAACAAAAACACCGCTACAGCGGGTTGTTGCGGATCAAGAGCTACCA  
 ACTCTTTCCGAGGTAACTGGCTCAGCAGAGGGAGATAACCAAAACTGTCCTCTA  
 GTGTAGCCGTAGTTAGGCCACCACTCAGAAGACTCTGTAGCACCAGGCCATACACCTCGCT  
 CTGCTAATCCTGTTACCGACTGGCTGCGAGTGGCGATAAGTCGTGTTACCGGGTTG  
 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGGTTG  
 ACACAGCCCAGCTGGAGCGAACGACCTACACCGAATGAGATACCTACAGCGTGAGCAT-

FIGURE 95B

224/240

TGAGAAAGGCCACGCTTCCCGAAGGGAGAAAGGCCAACGGTATCCGTAAGGGCAGG  
 GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCCTGGTATCTTATAGT  
 CCTGTGGGTTTGCACCTCTGACTTGAGCGTCGATTGATGCTCGTCAGGGGG  
 CCGAGCTATGGAAAAGCCAGCAACGCCCTTACGGTCTGCCCTTGCTGG  
 CCTTTGCTCACATGTTCTGCCTATCCCCTGATACTGTGGATAACCGTATTAC  
 GCCTTGAGTGAAGCTGATACCGCTGCCAGGCCAACGACCGAGCGAGTCAGTG  
 AGCGAGGAAGCGGAAGAGGCCAATACGAAACGCCCTCCCCGCCGTTGGCGATT  
 CATTAAATGCGACTGGCACAGGTTCCGACTGGAAAGCGGGAGTGAGCGAACGCA  
 ATTAATGAGTTACCTCATTAAGGCCAACGGCTTACACTTATGCTTCCGGCT  
 CCTATGTTGAGCTGAGCGGATAACAATTACACAGGAAACAGCTATGACCAT  
 GATTAGCCAAGCTCGGAATTAAACCCCTACTAAAGGAAACAAAGCTGGTACCGGGCC  
 CCCCTCGAGATCCGGGATCGAAGAAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATG  
 AAGGAAAAGACAAATAAGGTCGAACGAAAAAATAAGTGAAGAGTGTGATATGATG  
 TATTTGGCTTGCAGGGGAAACAGAGTTACGCAATTGACAATCATGCTGACTCT  
 GTGGCGGACCCCGCTCTTCCGGCGATAACGCTGGCGTAGGGCTGTGCCGGC  
 GGAGTTTGTGCCCTGCATTTCAGGTTACCTCGCTAACGGCGAGATTGGAGA  
 AGCAATAAGAATGCCGTTGGGTTCCGATGATGACGACCACGACAACGGTGTCTT  
 TTAAGTTGCCGAAAGAACCTGAGTGATTTGCAACATGAGTATACTAGAAGAATGAGCCA  
 AGACTTGCAGACCGAGTTGCCGGTGTGCGAACAAATAGAGCGACCATGACCTTGAAG  
 GTGAGACCGCATAACCGCTAGAGTACTTGAAGAGGAAACAGCAATAGGTTGCTACCA  
 GTATAAATAGACAGGTACATACAACACTGGAAATGGTTGCTGTTGAGTACGCTTCAA  
 TTCATTGGGTGCACTTTATTATGTTACAATATGAAAGGAACTTACACTTCTCTA  
 TGACACATATAATTAAATTAAAGTCCAATGCTAGTAGAGAACGGGGTAACACCCCTCCGCG  
 TCTTCCGATTTCCTAAACCGTGGAAATTTCGGATATCCTTGTGTTCCGGG  
 TGTACAATATGGACTCCTTTCTGGCAACCAAACCCATACATCGGATTCTATAAT  
 ACCTCGTTGGCTCCCTAACATGTAGGTGGCGAGGGAGATATACAATAGAACAGATA  
 CCAGACAAGACATAATGGCTAAACAGACTACACCAATTACACTGCCTCATGATGGTG  
 GTACATAACGAACTAATACTGTAGCCTAGACTTGATAGCCATCATATCGAAGTTTC  
 ACTACCTTTCTCATTTGCCATCTTGAAGTAATAATAGGCGATGCAACTCTTTTC  
 TTTTTTTCTCTCTCCCGTTGTCTCACCATATCCGAATGACAAAAAAA  
 ATGATGGAAGACACTAAAGGAAAAATTAAAGCACAAAGACAGCACCAACAGATGCGTTG  
 TTCCAGAGCTGATGAGGGTATCTCGAACACACGAAACTTTCTTCCTTCATTACAG  
 CACACTACTCTAAATGAGCAACGGTATACGCCCTCCAGTTACTGAATTGAAA  
 TAAAAAAAGTTGCCGTTGCTATCAAGTATAATAGACCTGCAATTATTAATCTT  
 TTCCCGTCAATTGTTCTCGTCCCTTCTCCTGTTCTTTCTGCACAATATTCA  
 AGCTATACCAAGCATAACATCAACTCCAAGCTTATGCCAAGAAGAAGCGGAAGGTCTCG  
 AGCGCGCCAATTAAATCAAAGTGGAAATTGCTGATAGCTCATTGCTCTTCACTTCA  
 ACTAACAGTAGCAACGGTCCGAACCTCATAACAACACTCAAACAAATTCTCAAGCGCTTCA  
 CAACCAATTGCCCTCTAACGTTCATGATAACTCATGAATAATGAAATACGGCTAGT  
 AAAATTGATGATGGAATAATTCAAACACTGTACCTGGTGGACGGACAAACTGCG  
 TATAACCGTTGGAATCACTACAGGGATGTTAATACCAACTACAATGGATGATGATAT  
 AACTATCTATTGATGATGAGATAACCCACAAACCAAAAAAGAGGGTGGGTCGAAT  
 CAAACAAGTTGTACAAAAAAGCTGAACGAGAAACGTAATGATATAATATCAATATA  
 TTAAATTAGATTGCTATAAAAACAGACTACATAATACTGTAAAACACAATATCCAG  
 TCACTATGGCGCCGCTAAGTGGCAGCATACCCGACGCACTTGCCTGCCAATAATAC  
 CTGTACGGAAGATCACTCGCAGAATAATAATCCTGGTGTCCCTGTTGATACCGGGA  
 AGCCCTGGCCAACCTTGGCAGAAATGAGACGTTGATCGGCACGTAAGAGGTTCCA  
 TTCACCATATGAAATAAGATCACTACCGGGCGTATTTTGAGTTATCGAGATTTCAG  
 GAGCTAAGGAAGCTAAAGGAGAAAAATCACTGGATATACCACCGTTGATATATCCC  
 AATGGCATTGAAAGAACATTGAGGATTCACTGAGCTCAATGTACCTATAACC  
 AGACCGTTCACTGGATATTACGCCCTTTAAAGACCGTAAAGAAAAATAAGCACAAGT  
 TTTATCCGCCCTTATTACACATTCTGCCGCTGATGAATGCTCATCGGAATTCCGTA  
 TGGCAATGAAAGACGGTGAGCTGGTATGGGATAGTGGTACCCCTGTTACACCGTT  
 TCCATGAGCAAACGTTTACCGCTCTGGAGTGAATACACGACGATTCCGG  
 AGTTTCTACACATATTGCAAGATGTGGCGTGTACGGTGAACCTGCCATTTC  
 CTAAGGGTTATTGAGAATATGTTTGTCTCAGCCAATCCCTGGGTGAGTTACCA  
 GTTTGATTAAACGTGCCAATATGGACAACCTTCTGCCCOGTTTACCATGGCA  
 AATATTATACGCAAGGCACAAGGTGCTGATGCCGCTGGCGATTCAAGGTTCATGCG-

Figure 95c

TCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGT  
GGCAGGGCGGGCGTAATCTAGAGGATCCGGTTACTAAAGCCAGATAACAGTATGCGT  
ATTTGCGCCTGATTTTGCCTATAAGAATATACTGATATGTATACCCGAAGTATGT  
CAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGCTATCA  
GTTGCTCAAGGCATATATGATGTCATATCTCCGTCTGTTAAGCACAACCAGTCAGAAT  
GAAGCCCCTCGTCTCGTGCCTGAAACGCGAAAATCAGGAAGGGATGGCTGAG  
GTCGCCGGTTATTGAAATGAACGGCTTTTGCTGACGAGAACAGGGACTGGTGAAT  
GCAGTTAAGGTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTGTGGATGTACA  
GAGTGATATTATTGACACGCCGGCGACGGATGGTATCCCCCTGCCAGTGCACGTCT  
GCTGTCAGATAAAAGTCTCCCGTGAACCTTACCCGGTGGTGCATATCGGGATGAAAGCTG  
GCGCATGATGACCACCGATATGCCAGTGTGCCGGTCTCCGTTATCGGGAAAGAAGTGGC  
TGATCTCAGCCACCGCGAAAATGACATCAAAACGCCATTAAACCTGATGTTCTGGGAAT  
ATAAAATGTCAGGCTCCGTTATACACAGCCAGTCTCAGGTCACCAGTGAATGGATAT  
GTTGTTTACAGTATTATGTTGTTGAAATCTAATTAAATATTGA  
TATTATATCATTTCAGTTCTCGTTCTGTTGACAAAGTGGTTGATGGCCGC  
TAAGTAAGTAAGACGTCGAGCTCCCTATAGTGAGTCGTTACACTGCCGTCGTTTAC  
AACGTCGTGACTGGGAAAACACCGGTGAGCTTAAGTAAGTAACGCCGCCACCGCGGTG  
GAGCTTGGACTCTCGCCAGAGGTTGGTCAAGTCTCCAATCAAGGTTGTCGGCTTGT  
CTACCTGCCAGAAATTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTG  
TTGACACTCTAAATAAGCGAATTCTTATGATTATGATTGTTATTAAATAAGTTA  
TAAAAAAAATAAGTGATACAAATTAAAGTGACTCTTAGGTTAAAACGAAAATTCT  
TGTTCTTGAGTAACCTTCTGTAGGTCAAGGTTCTCAGGTATAGCATGAGGTG  
CTCTTATTGACCACACCTCACCGCATGCCAGCAAATGCTGCAAATGCTCCCCATT  
TCACCCAATTGAGATATGTAACCTCAGCAATGAGGTTGATGAATCTCGGTGTTGATT  
ATGTCCTCAGAGGACAATACTGTTGTAATCGTTCTCACCGATCCGATCAGCGA  
AATTGTAACGTTAATATTGTTAAAATCGCGTTAAATATTGTTAAATCAGCTCATT  
TTTAACCAATAGGCCGAAATCGGAAAATCCCTTATAAATCAAAGAATAGACCGAGAT  
AGGGTTGAGTGTGTTCCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA  
CGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGCCACTACGTGAACCATCACCTA  
ATCAAGTTTTGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCC  
CCGATTAGAGCTTGACGGGAAAGCCGGCAACGTGGCAGAAAGGAAGGGAAAGAACG  
GAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGAGCGGTACGCTGCGCTAACCAAC  
ACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTGCCATTCACTGCA

FIGURE 95D

226/240

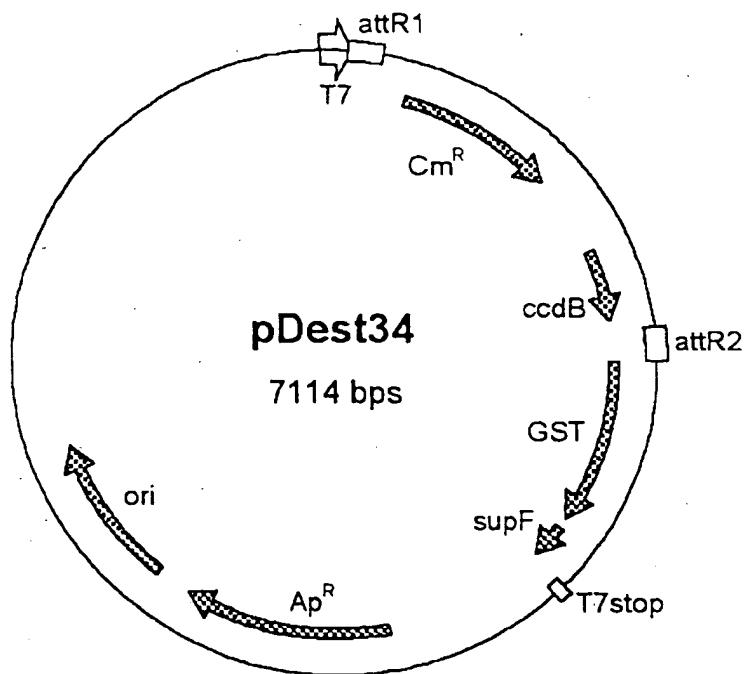


FIGURE 96A

227/240

pDEST34 7114 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
195..71	attR1
304..963	CmR
1305..1610	ccdB
1651..1775	attR2
1780..2472	GST
2675..2720	T7stop
3334..4194	ampR
4343..4982	ori

ATCGAGATCTCGATCCCGGAAATTAAACGACTCACTATAGGGAGACCACAACGGTTTC  
 CCTCTAGATCACAAAGTTGTACAAAAAAGCTGAACGAGAAACGTAATGATATAAATAT  
 CAATATATTAAATTAGATTTCGATAAAAAACAGACTACATAACTGTAAAACACAACA  
 TATCCAGTCACTATGGCGGCCGCATTAGGCACCCCAGGCTTACACTTATGCTTCCGGC  
 TCGTATAATGTGGATTGAGTTAGGATCCGGCAGAGATTTCAGGAGCTAAGGAAGCT  
 AAAATGGAGAAAAAAATCACTGGATATACCAACCGTTGATATATCCAATGGCATCGTAA  
 GAACATTTGAGGCATTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTG  
 GATATTACGGCTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTATCCGGCTTT  
 ATTACACATTCTGCCGCTGATGAATGCTCATCCGAATTCCGTATGGCAATGAAAGAC  
 GGTGAGCTGGTGATATGGGATAGTGGTACCCCTGTTACACCGTTTCCATGAGCAAAC  
 GAAACGTTTCATCGCTCTGGAGTGAATACCAACGACGATTCCGGCAGTTCTACACATA  
 TATTGCAAGATGTGGCTGTTACGGTAAACCTGGCTTATTCCTAAAGGGTTTATT  
 GAGAATATGTTTCGTCAGCCAATCCCTGGGTGAGTTTACCAAGTTGATTTAAAC  
 GTGGCCAATATGGACAACCTCTTCGCCCGGTTTACCATGGGAAATATTATACGCAA  
 GGCGACAAGGTGCTGATGCCGCTGGCATTAGGTTCATCATGCCGCTGTGATGGCTTC  
 CATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCG  
 TAAACCGTGGATCCGGCTACTAAAGCCAGATAACAGTATGCGTATTGCGCGCTGAT  
 TTTGCGGTATAAGAATATATACTGATATGTATACCGAAGTATGTCAAAAGAGGTGTG  
 CTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCAT  
 ATATGATGTCAATATCTCCGGTCTGGTAAGCACAACCATGCGAATGAAGCCCCTCGTCT  
 GCGTGCCGAACGCTGGAAAGCGGAAATCAGGAAGGGATGGCTGAGGTCGCCGGTTAT  
 TGAAATGAACGGCTTTTGCTGACGAGAACAGGGACTGGTAAATGCAAGTTAAGGTTT  
 ACACCTATAAAAGAGAGAGCCATTACGTCTGTTGTGATGTACAGAGTGTATTATTG  
 ACACGCCGGCGACGGATGGTATCCCCCTGGCAGTGCACGTCTGTCAGATAAAG  
 TCTCCCGTGAACCTTACCCGGTGGTCATATCGGGGATGAAAGCTGGCGATGACCA  
 CCGATACTGGCAGTGTGCCGCTCCGTTATCGGGGAAAGAAGTGGCTGATCTCAGCCACC  
 GCGAAAATGACATCAAAACGCCATTACGTGTTCTGGGAAATATAATGTCAGGCT  
 CCCTATACACAGCCAGTCTGCAGGTCGACCATAGTGACTGGATATGTTTACAG  
 TATTATGATGCTGTTTATGAAAATCTAATTAAATATTGATATTATATCATTT  
 TACGTTCTCGTTCACTTCTGACAAAGGGTGTATTGTCCTTACTAGGTTAT  
 TGGAAAATTAAAGGGCCTTGTGCAACCCACTCGACTTCTTTGGAAATATCTGAAAGAAAA  
 TATGAAGAGCATTGATGAGCGCGATGAAGGTGATAATGGCAAACAAAAAGTTGAA  
 TTGGGTTGGAGTTCCCAATCTCCTTATTATATTGATGGTGTGTTAAATTAAACACAG  
 TCTATGCCATCATCGTTATATAGCTGACAAGCACAACATGTTGGGTGGTGTCCAAAA  
 GAGCGTGCAGAGATTCAATGCTGAAGGAGCGGTTTGGGATATTAGATACGGTGTTCG  
 AGAATTGCAATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTCTTAGCAAGCTACCT  
 GAAATGCTGAAAATGTTGCAAGATCGTTATGTCATAAAACATTTAAATGGTGTACAT  
 GTAACCCATCTGACTTCATGTTGATGACGCTCTGATGTTGTTTATACATGGACCCA  
 ATGTCCTGGATGCGTTCCAAAATTAGTTGTTTAAACGATTGAAAGCTATCCCA  
 CAAATTGATAAGTACTGAAATCCAGCAAGTATAGCATGGCCTTGCAGGGCTGGCAA  
 GCCACGTTGGTGGCGACCATCCTCCAAAATCGGATCTGGTCTCCGGTCCATGGGGA  
 TCCGGCTGCTAACAAAGGCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCGCTT  
 CCCGATAAGGGAGCAGGCCAGTAAAGCATTACCGTGGTGGGGTTCCGAGCAGGCCAAA  
 GGGAGCAGACTCTAAATGCGCTCATCGACTTCGAAGGTTCGAATCCTCCCCCACAC  
 CATCACTTCAAAAGTGAATTGCGTGAAGCAATAACTAGCATAACCCCTGGGGCTCTAA-

FIGURE 96B

ACGGGTCTGAGGGGTTTTGCTGAAAGGAGGAACATACTCCGGATATCCACAGGACGG  
 GTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTG  
 GGCAGCGGCCAACAGCGGTCGGACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCA  
 ACGCATATAGCGCTAGCAGCACGCCATAGTGAATGGCGATGCTGCGAATGGACGATAT  
 CCCGCAAGAGGCCCGCAGTACCGGCATAACAAAGCCTATGCCCTACAGCATCCAGGGTGA  
 CGGTGCCGAGGATGACGATGAGCGCATTGTTAGATTTCATACACGGTGCCTGACTGCGTT  
 AGCAATTAACTGTGATAAAACTACCGCATTAAAGCTTATCGATGATAAGCTGCAAACAT  
 GAGAATTCTGAAGACGAAAGGGCTCGTGATACGCCATTTTTATAGGTTAATGTCATG  
 ATAATAATGGTTCTTAGACGTAGGTGCACTTTCGGGAAAATGTGCCGGAAACCCCT  
 ATTTGTTTATTCTAAATACATTCAAATATGTATCCGCTATGAGACAATAACCCCTGA  
 TAAATGCTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTCCGTGCGCC  
 CTTATTCCCTTTTGCGGCATTTCGCTTCTGTTGCTCACCCAGAAACGCTGGTG  
 AAAGTAAAGATGCTGAAGATCAGTTGGTGCACGAGTGGTTACATCGAACTGGATCTC  
 AACAGCGTAAGATCCTGAGAGTTTCGCCCGAAGAACGTTTCCAATGATGAGCACT  
 TTAAAGTTCTGCTATGTGGCGGTATTATCCGTGTTGACGCCGGCAAGAGCAACTC  
 GGTCGCCGATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAG  
 CATCTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCCTACATGAGTGAT  
 AACACTGCGGCCACTTACTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTT  
 TTGACACACATGGGGATCATGTAACTCGCCTGATCGTTGGAAACCGAGCTGAATGAA  
 GCCATACCAACGACGAGCGTACACCGATGCCGCAGCAATGGCAACACGTTGCGC  
 AAACATTAACGCGAACTACTTACTCTAGCTCCGGCAACAATTAAAGACTGGATG  
 GAGGCAGATAAGGTTGAGGACCACTCTGCGCTCGGCCCTCCGGCTGGCTGGTTATT  
 GCTGATAAAATCTGGAGCCGGTGGCTGAGCTGGTCTCGGGTATCATTGAGCAGTGGGCA  
 GATGGTAAGCCCTCCGTATCGTAGTTACACGACGGGAGTCAGGCAACTATGGAT  
 GAACGAAATAGACAGATCGTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCA  
 GACCAAGTTACTCATATATACTTAGATTGATTAAACTCATTAAATTAAAGG  
 ATCTAGGTGAAGATCCTTTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTCC  
 TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTT  
 CTGCGCTAATCTGCTGTTGCAAACAAAAACACCGCTACAGCGGTGGTTGTTG  
 CCGGATCAAGAGCTACCAACTCTTTCGAAGGTAACGGCTTCAGCAGAGCGCAGATA  
 CCAAATACTGCTCTAGTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGAGCA  
 CCGCCTACATACCTCGCTGCTAACCTGTTACAGTGGCTGCTGCCAGTGGCGATAAG  
 TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGC  
 TGAACGGGGGTTCTGACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGA  
 TACCTACAGCGTGAGCTATGAGAAAGGCCACGCTTCCGAAGGGAGAAAGGCCGACAGG  
 TATCCGTAAGCAGCAGGGTGGAAACAGGAGAGCGCAGGAGGGAGCTCCAGGGGAAAC  
 GCCTGGTATCTTATAGTCTGCGGGTTGCCCCCTGACTGAGCGTCACTGAGCTGATTTG  
 TGATGCTCGTCAAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGCCCTTTACGG  
 TTCCCTGCCCTTTGCTGCCCTTGCTCACATGTTCTTCTGCGTTATCCCTGATTCT  
 GTGGATAACCGTATTACCGCCTTGAGTGGCTGATACCGCTGCCAGCCGAACGACC  
 GAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCTGATGCCGTATTCTCCT  
 ACGCATCTGCGGTATTCACACCGCATATATGGCACTCTCAGTACAATCTGCTCTG  
 ATGCCGATAGTTAAGCCAGTATAACACTCCGCTATCGCTACGTGACTGGTCTGGCTGC  
 GCCCCGACACCCGCCAACACCCGCTGACGCCCTGACGGCTGCTGCCGGCATC  
 CGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTCAGAGGTTTACCGTC  
 ATCACCGAAACGCGCAGGCAGCTGGTAAAGCTCATCGCTGGTGTGAAGCGATT  
 ACAGATGCTGCCGTTACCGCCTCAGCTCGTTGAGTTCTCCAGAAGCGTTAATGT  
 CTGGCTCTGATAAAAGCGGCCATGTTAAGGGCGTTTTCTGTTGGTCACTGATGC  
 CTCCGTTAAGGGGATTCTGTTATGGGGTAATGATACCGATGAAACGAGAGAGGAT  
 GCTCACGATAACGGTTACTGATGATGAAACATGCCGGTACTGGAACGTTGTGAGGGTAA  
 ACAACTGGCGGTATGGATGCCGGGACCAAGAGAAAATCACTCAGGGTCAATGCCAGCG  
 CTTCGTTAACAGATGTTAGGTGTTCCACAGGGTAGCCAGCAGCATCTGCGATGCAGAT  
 CCGGAACATAATGGTCAGGGCGCTGACTTCCGCTTCCAGACTTACGAAACACGGAA  
 ACCGAAGACCATTCACTGTTGCTCAGGTGCGAGACGTTTGCGAGCAGCAGTCGCTTCA  
 CGTTCGCTCGCGTATCGGTGATTCATCTGCTAACAGTAAGGCAACCCCGCCAGCCTAG  
 CGGGGTCTCAACGACAGGAGCACGATCATGCCACCGTGGCCAGGACCCAACGCTGCC  
 CGAGATGCGCCGCGTGCAGGGCTGCTGGAGATGGCGGACGGCATGGATATGTTCTGCCAAGG  
 GTTGGTTGCGCATTACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTGGAGTGGT-

FIGURE 96C

229/260

GAATCCGTTAGCGAGGTGCCGCCGGCTTCATTCAGGTCGAGGTGCCCGGCTCCATGCA  
CCGCACGCAACCGGGGAGGCAGACAAGGTATAAGGGCGCGCTACAATCCATGCCAAC  
CCGTTCCATGTGCTCGCCGAGGCGGATAAAATCGCCGTGACGATCAGCGGTCCAGTGATC  
GAAGTTAGGCTGGTAAGAGCCCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTATCT  
ACCTGCCTGGACAGCATGGCCTGCAACCGGGCATCCCGATGCCGCCGAAAGCGAGAAGA  
ATCATATAATGGGGAAAGGCCATCCAGCCTCGCGTGCAGCACGCAAGACGTAGCCCAGC  
GCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTCTCGCCGAAACGTTGGTGGCGGG  
CCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCC  
ATCATCGTCGCGCTCCAGCGAAAGCGGTCCCTGCCGAAAATGACCCAGAGCGCTGCCGC  
ACCTGTCCCTACGAGTTGCATGATAAAAGAACAGTCATAAGTGCAGCGACGATAGTCATG  
CCCCCGCGCCACCAGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTGATCG  
ACGCTCTCCCTTATGCAGCTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTT  
GAGCACCGCCGCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGC  
CACGGGGCCTGCCACCATACCCACGCCGAAACAAAGCGCTCATGAGCCGAAGTGGCGAGC  
CCGATCTTCCCCATCGGTGATGTCGGCGATATAAGGCGCCAGCAACCGCACCTGTGGCGCC  
GGTGTGCGGCCACGATGCGTCCGGCGTAGAGG

FIGURE 9(a)

230/240

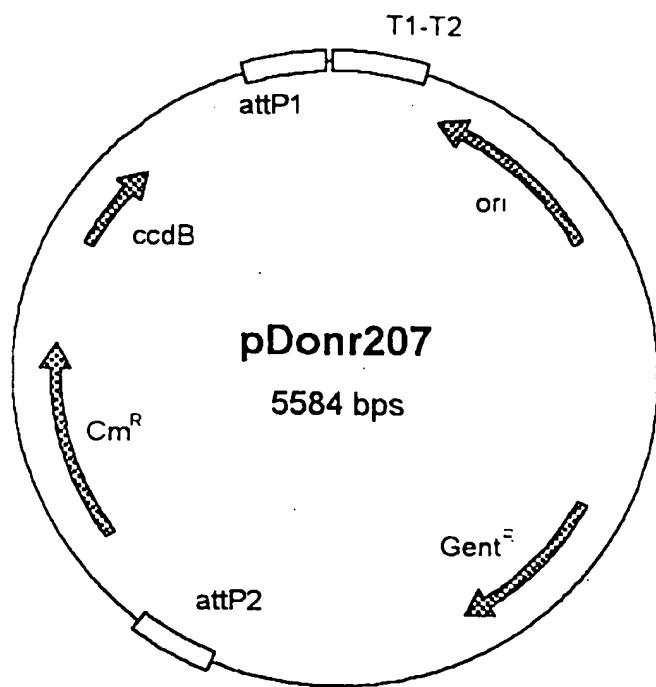


FIGURE 97A

pDONR207 5584 bp

GCGAGAGTAGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGGAAAGACTGGGC  
 CTTTCGTTTATCTGTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGG  
 AGCGGATTGAACTGTGAAGCAACGGCCGGAGGGTGGCGGGCAGGACGCCATA  
 AACTGCCAGGCATCAAACTAAGCAGAAGGCATCCTGACGGATGGCTTTTGCGTTCT  
 ACAAAACTCTTCCTGGCTAGCGTAATACGGTTATCCACAGAAATCAGGGATAACGCAGGA  
 AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGTTGCTG  
 GCGTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCAGCCTCAAGTCAG  
 AGGTGGCGAAACCCGACAGGACTATAAGATAACCAGGCGTTCCCCCTGGAAGCTCCCTC  
 GTGCCTCTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCG  
 GGAAGCGTGGCGTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGTGTAGGTCGTT  
 CGCTCCAAGCTGGCTGTGACGAAACCCCGTTGACGCCGACCGCTGCGCCTTATCC  
 GGTAACTATCGTCTTGAATGCTAACCCGGTAAGACACGACTATGCCACTGGCAGCAGCC  
 ACTGGTAACAGGATTAGCAGAGCAGGGTATGTAAGGCGGTGCTACAGAGTTCTGAAGTGG  
 TGGCCTAATCAGGCTACACTAGAAGGACAGTATTGGTATCTGCCTCTGCTGAAGCCA  
 GTTACCTCGGAAAAGAGTTGGTAGCTTGTGATCCGCAACAAACACCACCGCTGGTAGC  
 GGTGGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGAT  
 CCTTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATT  
 TTGGTCATGAGCTGCGCCGTCAGTCAGCGTAATGCTCTGCCAGTGTACAACC  
 AATTAACCAATTCTGATTAGAAAACATCGAGCATCAAATGAAACTGCAATTATTCA  
 TATCAGGATTATCAATACCATATTTGAAAAAGCCGTTCTGTAATGAAGGAGAAAAC  
 CACCGAGGCAGTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCCATTCCGACTCGTC  
 CAACATCAATACAACCTATTAGTAGCCAACCAACTAGAAACTATAGCTAGAGTCCTGGCGA  
 ACAAAAGATGCTGCCCTCCAGAAAACCGAGGATGCGAACCACTTCATCCGGGTCAGCA  
 CCACCGCAAGCGCCCGCAGGCCAGGGTCTCCGATCTCTGAAAGCCAGGGCAGATCCG  
 TGCACAGCACCTGCGTAGAAGAACAGCAAGGGCGCAATGCCCTGACGATGCGTGGAGA  
 CCGAAACCTTGCCTCGTCGCCAGCAGGACAGAAATGCCCTGACTTCGCTGCTGCCA  
 AGGTTGCCGGGTGACGCACACCGTGGAAACGGATGAAGGACAGAACCCAGTTGACATAAG  
 CCTGTTGGTCTGTAACACTGTAATGCAAGTAGCGTATGCCCTCACGAACTGGTCCAGAA  
 CCTTGACCGAACCGCAGCGGTGTAACGGCGCAGTGGCGTTTATGGCTTATGACT  
 GTTTTTTGTACAGTCTATGCCCTGGCATCCAAGCAGCAAGCGCTTACGCCGTGGTC  
 GATGTTGATGTTATGGAGCAGCAACGATGTTACGCGAGCAACGATGTTACGAGCAG  
 GGCAGTCGCCCTAAAACAAAGTTAGGGCTCAAGTATGGCATATTGCACTATGTAGG  
 CTCGGCCCTGACCAAGTCAAATCCATGCCCTGCTTGTATCTTCTGGTGTGAGTT  
 GGAGACGTAGCCACCTACTCCAAACATCAGCCGACTCCGATTACCTGGGAACTTGCTC  
 CGTAGTAAGACATTATCGCCTGCTGCCACCAAGAAGCGTTGTTGGCCTCTC  
 GCGGCTACGTTCTGCCAGGTTGAGCAGCCCGTAGTGGATCTATATGATCTC  
 GCAGTCGCCGGCAGCACCGAGGCAGGGCATTGCCACCGCCTCATCAATCTCTCAAG  
 CATGAGGCCAACCGCCTGGTCTATGTGATCTACGTGCAAGCAGATTACGGTACGAT  
 CCCGAGTGGCTCTATACAAAGTGGGATCACGGAAAGAAGTGTGACTTTGATATC  
 GACCCAAAGTACCGCACCTAACATTGTTCAAGCGAGATGGCTCCCGGCTAATT  
 CCCCTGTCAAAATAAGTTATCAAGTGAGAAATCACCAGAGTACGACTGAATCCGG  
 TGAGAATGGAAAAGTTATGCAATTCTTCCAGACTTGTCAACAGGCCAGCCATTACG  
 CTCGTCATCAAAACTCGCCTAACCAACCGTATTCTGATGCCCTGAGC  
 GAGACGAAATACCGATGCTGTTAAAGGACAATTACAAACAGGAATGCAACCC  
 GCGCAGGAACACTGCCAGCGCATCAACAAATTTCACCTGAAATCAGGATATTCTCTAA  
 TACCTGGAATGCTGTTTCCGGGATCGCAGTGGTAGAGTAACCATGCACTCAGGAGT  
 ACGGATAAAATGCTTGATGGCTGGAAAGAGGCAAAATTCCGTCAGCCAGTTAGTCTGAC  
 CATCTCATCTGTAACATCATTGCCAACGCTACCTTGCCATGTTGAGAAACAACTCTGG  
 CGCATCGGGCTTCCCATAACAGCGATAGATTGTCGCACCTGATTGCCGACATTATCGCG  
 AGCCCCATTATACCCATAAAATCAGCATCCATGTTGAAATTAAATGCCGCTCGACGT  
 TTCCCGTTGAATATGGCTCATACACACCCCTGTATTACTGTTTATGTAAGCAGACAGTT  
 TATTGTTGATGATGATATTTTATCTGTCATGAACTCAGAGATTGAGACAC  
 GGGCCAGAGCTGAGCTGGATGGCAAAATAATGATTGACTGATGACCTGTT  
 CGTTGCAACAAATTGATAAGCAATGCTTCTTATAATGCCAACTTGTACAAGAAAGCTG  
 AACGAGAAACGTAACATGATATAAATCAATATATTAAATTAGATTGATGATGAA  
 AGACTACATAACTGTAACACACATATCCAGTCACATGAACTACTTAGATG-

FIGURE 975

232/240

GTATTAGTGACCTGAGTCGACTAAGTTGGCAGCATACCCGACGCACTTGCGCCGAAT  
AAATACCTGTGACGGAAGATCACTTCCAGAATAAAATAATCCTGGTGTCCCTGTTGATA  
CCGGGAAGCCCTGGGCCAACTTGGCAAAATGAGACGTTGATCGGCACGTAAGAGGTTC  
CAACTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTGAGTTATCGAGATT  
TTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAAACTGGATATACCAACCGTTGATAT  
ATCCCAATGGCATCGTAAAGAACATTGAGGCATTCAGTCAGTGCTCAATGTACCTA  
TAACCAGACCGTTAGCTGGATATTACGGCCTTTAAAGACCGTAAAGAAAAATAAGCA  
CAAGTTTATCCGGCCTTATTACACATTCTGCCCGCTGATGAATGCTCATCCGGAATT  
CCGTATGGCAATGAAAGACGGTGAGCTGGTATGGGATAGTGTTCACCTTGTACAC  
CGTTTCCATGAGCAAACGTTTCATCGCTCTGGAGTGAATACCACGACGATTT  
CCGGCAGTTCTACACATATTCGCAAGATGTGGCGTTACGGTAAAACCTGGCCTA  
TTTCCCTAAAGGGTTATTGAGAATATGTTTCTGCTCAGCCAATCCCTGGGTGAGTTT  
CACCAAGTTTGTATTAAACGTGCCAATATGGACAACCTCTCGCCCCGTTTACCAT  
GGCAAAATATTACGCAAGGCAGAACAGGTGCTGATGCCGCTGGCATTAGGTCATCA  
TGCCGCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGA  
TGAGTGGCAGGGCGGGCGTAATCGCGTGGATCCGCTTACTAAAAGCCAGATAACAGTA  
TGCATTTGCGCCTGATTTCGCGTATAAGAATATACTGATATGTATAACCGAAG  
TATGTCAAAAGAGGTGCTGATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGC  
TATCAGTTGCTCAAGGCATATATGATGTCATATCTCCGGTCTGGTAAGCACAACCATGC  
AGAATGAAGCCCCTCGTCTCGTGCCTGGAAACGCTGGAAAAGCGAAAATCAGGAAGGGATGG  
CTGAGGTGCCCCGTTATTGAAATGAACGGCTTTGCTGACGAGAACAGGGACTGGT  
GAAATGCAGTTAAGGTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTGTGGAT  
GTACAGAGTGATATTATTGACACGCCGGCGACGGATGGTATCCCCCTGGCCAGTGCA  
CGTCTGCTGTCAGATAAGCTCCCGTGAACCTTACCGGGTGTGCATATCGGGGATGAA  
AGCTGGCGCATGATGACCAACGATATGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAA  
GTGGCTGATCTCAGCCACCGAAAATGACATCAAAACGCCATTAACCTGATGTTCTGG  
GGAATATAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTGATACAGTAGAAAT  
TACAGAAACTTATCACGTTAGTAAGTATAGAGGCTGAAAATCCAGATGAAGCCGAACG  
ACTTGTAAGAGAAAAGTATAAGAGTTGTGAAATTGTTCTGATGCAAGATGATTTCAGGA  
CTATGACACTAGCGTATATGAATAGGTAGATGTTTATTTGTCACACAAAAAGAGGC  
TCGCACCTTTCTTCTTATGATTAAATACGGCATTGAGGACAATAGCGAG  
TAGGCTGGATACGACGATCCGTTGAGAAGAACATTGGAAGGCTGCGGTGACTAAG  
TTGGCAGCATACCCGAAGAACATTGGAAGGCTGCGGTGACTACAGGTCACTAATAC  
CATCTAAGTAGTTGATTCAAGTGACTGGATATGTTGTTTACAGTATTATGAGTC  
GTTTTTATGCAAAATCTAATTAAATATTGATATTATCATTACGTTCTCGTT  
CAGCTTTTGTACAAAGTGGCATTATAAAAAGCATTGCTCATCAATTGTTGCAACG  
AACAGGTCACTATCAGTCAAAATAAAATCATTATTGGGGCCCGAGATCCCATGCTAGCGT  
TAAC

FIGURE 97C

233/240

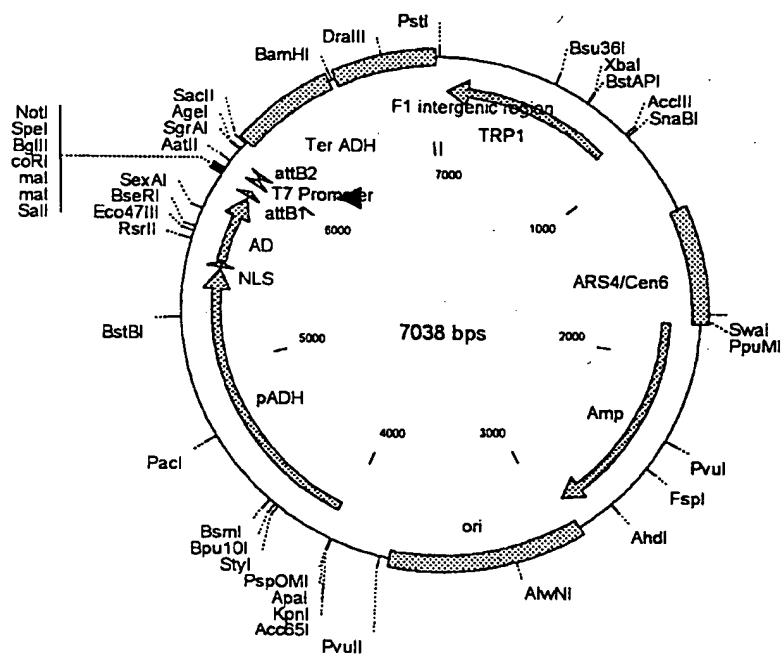
**pMAB85**

FIGURE 98A

234/240

pMAB85 7038 bp

GCCTTACGCATCTGTGGGTATTCACACCGCAGGCAAGTCACAAACAATACTTAAATA  
 AATACTACTCAGTAATAACCTATTCTTAGCATTTGACGAAATTGCTATTGTTAG  
 AGTCTTTACACCATTGCTCCACACCTCCGTTACATCAACACCAATAACGCCATT  
 ATCTAACGCGCATACCAACATTCTGGCGTCAGTCCACAGCTAACATAAAATGTAAGC  
 TTTCGGGGCTCTTGCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCAC  
 CTGCCCACCTGCTCTGAATCAAACAAGGGATAAACGAATGAGGTTCTGTGAAGCTG  
 CACTGAGTAGTATGTTGAGTCAGTTGGAAATACGAGTCAGTTAATAACTGGAAACCGA  
 GGAACCTTGGTATTCTGCCACGACTCATCTCCATGAGTTGGACGATATCAATGCCGT  
 AATCATTGACCAAGGCCAAACATCCTCTTAGGTTGATTACGAAACACGCCAACCAAGT  
 ATTCGGAGTGCTGAACATATTATGCTTTACAAGACTTGAATTTCTTGCAA  
 TAACCGGGTCAATTGTTCTTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCAT  
 CGGAATCTAGAGCACATTCTGGCCCTCTGTGCTCTGCAAGCCGAAACTTCACCAATG  
 GACCAGAACTACCTGTGAAATTAAATAACAGACATACTCCAAGCTGCCCTTGCTGCTAA  
 TCACGTATACTCACGTGCTCAATAGTCACCAATGCCCTCCCTTGCCCTCCCTTT  
 TTTTCGACCGAATTAAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTG  
 ACGTAAGGTGACAAGCTATTGCAATAAGAATATCTTCACTACTGCCATCTGGCGC  
 ATAACGTCAAAGTACACATATAATTACGATGCTGCTATTAAATGCTTCTATATTATA  
 TATAGTAATGTCGTTATGGTGCACTCTAGTACAATCTGCTCTGATGCCATAGTTAA  
 GCCAGCCCCGACACCCGCCAACACCCGCTGACGCCCTGACGGGTTGTCTGCCCG  
 CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTCAC  
 CGTCATCACGAAACGCCGAGACGAAAGGGCCTCGTGTACGCCATTAGTTAGGTTA  
 ATGTCATGATAATAATGGTTCTTAGGACGGATCGCTGCCTGTAACACGCCCTC  
 GTATTTTAATGATGAAATAATTGGGAAATTACTCTGTTATTATTTATGTT  
 TGTATTGGATTAGAAAGTAAATAAGAAGGTTAGAGAGTACGGAATGAAGAAAAAA  
 AAATAACAAAGGTTAAAAAAATTCAACAAAAGCGTACTTACATATAATTAG  
 ACAAGAAAAGCAGATTAAATAGATATACTCGATTAACGATAAGAAAATGAAATCA  
 CAGGATTTCTGTTGCTTCTACACAGACAAGATGAAACAATTGGCATTAAACACT  
 GAGAGCAGGAAGAGCAAGATAAAAGGTTAGTATTGTTGGCATTCCCTAGAGTC  
 CATCTCGGAAACAAAACATTTTCTTAAATTCTTTTACTTCTATTAA  
 TTATATATTTATTTAAATTATAATTATTTATAGCACGTGATGAAAG  
 GACCCAGGTGGCACTTTGGGAAATGTGCGGAAACCCCTATTGTTATTCTAA  
 ATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAAATGCTCAATAATAT  
 TGAAAAGGAAGAGTATGAGTATTCAACATTCCGTTGCGCCCTATTCCCTTTGCG  
 GCATTTCGCTTCCGTGTTGCTACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAA  
 GATCAGTTGGGTGACGAGTGGTTACATGAACTGGATCTCAACAGCGGTAAAGATC  
 GAGAGTTTCGCCCGAAGAACGTTCCAAATGATGAGCACTTTAAAGTTCTGCTATGT  
 GGCGGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTCGCCGATACACTAT  
 TCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTACGGATGGCATG  
 ACAGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAACACTGCCAACTTA  
 CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTCAAAACATGGGGAT  
 CATGTAACTCGCCTTGATCGTTGGAAACGGAGCTGAATGAAGCCATACCAACGAG  
 CGTGACACCACCGATGCCGTAGCAATGGCAACAAACGTTGCGCAAACATTAACTGGGAA  
 CTACTACTCTAGCTCCCGCAACAATTAAAGACTGGATGGAGGGGAAAGTTGCA  
 GGACCACTCTGCGCTGCCCTCCGGCTGGTTATTGCTGATAAAATCTGGAGGCC  
 GGTGAGCGTGGTCTCGGGTATCATTGAGCACTGGGGCCAGATGGTAAGCCCTCCCGT  
 ATCGTAGTTATCTACACGACGGGAGTCAGGAAACTATGGATGAACGAAATAGACAGATC  
 GCTGAGATAGGTGCCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTACTCATAT  
 ATACTTTAGATTGATTAAAACCTCATTTAAATTAAAGGATCTAGGTGAAGATC  
 TTTGATAATCTCATGACCAAAACCTTAACGTGAGTTCTGTTCCACTGAGCGTCAGAC  
 CCCGTAGAAAAGATCAAAGGATCTCTTGAGATCCTTCTGCGCGTAATCTGCTGC  
 TTGCAAACAAAAACACCGCTACCGGGTGGTTGTTGCCGATCAAGAGCTACCA  
 ACTCTTTTCCGAAGGTAACGGCTCAGCAGAGCGCAGATACCAAATCTGCTCTA  
 GTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGAGCACCCTACATACCTCGCT  
 CTGCTAATCTGTTACCAAGTGGCTGCGAGTGGCGATAAGTCGTCTTACGGGTTG  
 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGTTCGTGC-

Figure 98B

235 / 240

ACACAGCCCAGCTGGAGCGAACGACCTACACCGAACTGAGATACTACAGCGTGAGCAT  
 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG  
 GTCGGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGGAAACGCCCTGGTATCTTATAGT  
 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTGTGATGCTCGTCAGGGGG  
 CCGAGCCTATGGAAAACGCCAGCAACGCGGCCCTTTACGGTTCTGGCCTTGTGG  
 CCTTTGCTCACATGTTCTCGCTTATCCCTGATTCTGTGAGATAACCGTATTACC  
 GCCTTGAGTGAGCTGATACCGCTGCCAGCGAACGACCGAGCGCAGCGAGTCAGTG  
 AGCGAGGAAGCGGAAGAGCGCCAATACGCAAACGCCCTCCCCCGCGTTGGCGATT  
 CATTAATGCACTGGCACGACAGGTTCCGACTGGAAAGCGGGCAGTGAGCGAACGCA  
 ATTAATGAGTTACCTCACTCATTAGGCACCCAGGTTACACTTATGCTTCCGGCT  
 CCTATGTTGAGTGGAAATTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCAT  
 GATTACGCCAACGCTCGGAATTAAACCCCACTAAAGGGAACAAAAGCTGGTACCGGGCCC  
 CCCCTCGAGATCGGGATCGAAGAAATGATGGAAATAGGAAATCAAGGAGCATG  
 AAGGAAAAGACAAATATAAGGGTCAAACGAAAAAATAAGTGAAGTGGTATGATG  
 TATTGGCTTGGCGCCGAAAAAACGAGTTACGCAATTGACAATCATGCTGACTCT  
 GTGGCGGACCGCGCTTGGCGCCCGGATAACGCTGGCGTGGCTGAGGCTGTGCCGGC  
 GGAGTTTGGCCCTGCATTTCAGGTTACCCCTGCGCTAACGGGGCAGATTGGAGA  
 AGCAATAAGAATGCCGTTGGGTTGCGATGACGACACGACAACACTGGTGTCAATT  
 TTAAGTTGCCGAAAGAACCTGAGTGCATTGCAACATGAGTATACTAGAAGAATGAGCCA  
 AGACTTGCAGACGCGAGTTGCCGTTGCGAACATAGAGCGACCATGACCTTGAAG  
 GTGAGACGCCATAACCGCTAGAGTACTTGAAGAGGAAACAGCAATAGGGTGTACCA  
 GTATAAAATAGACAGGTACATACAACACTGGAAATGGTTGTCTGTTGAGTACGCTTCAA  
 TTCATTGGGTGCACTTATTATGTTACAATATGGAAGGGAACTTACACTTCTCTTA  
 TGACACATATAATTAAATTAAGTCAATGCTAGTAGAGAACGGGGTAACACCCCTCGCGC  
 TCTTCCGATTTTCTAAACCGTGGAAATTTCGGATATCCTTGTGTTCCGG  
 TGTACAATATGGACTTCTCTTCTGGCAACCAACCCATACATGGATTCTATAAT  
 ACCTCGTTGGCTCCCTAACATGTAGGTGGCGAGGGAGATACAATAGAACAGATA  
 CCAGACAAGACATAATGGCTAAACAGACTACACCAATTACACTGCCCTATTGATGGT  
 GTACATAACGAACATAACTGTAGCCCTAGACTTGAAGCTCATCATATCGAAGTTTC  
 ACTACCTTTTCCATTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTC  
 TTTTTTTCTTCTCTCTCCCCGTTGTCACCATATCCGAATGACAAAAAAA  
 ATGATGGAAGACACTAAAGGAAAAATTAAAGCACAAGACAGCACCAACAGATGCGTTG  
 TTCCAGAGCTGATGAGGGTATCTCGAACACACGAAACTTTCTCTCATTCA  
 CACACTACTCTAAATGAGCAACGGTATACGCCCTCCAGTTACTGAATTGAAA  
 TAAAAAAAGTTGCCGTTGCTATCAAGTATAATAGACCTGCAATTATTAAATCTTTG  
 TTCCCTCGTCATTGTTCTGTTCCCTTCTCCTGTTCTGACAATATTCA  
 AGCTATACCAAGCATAACAATCAACTCCAGCTTATGCCAAGAAGAACGGAAAGGTCTG  
 AGCGCGCCAATTAAATCAAAGTGGAAATTGCTGATAGCTATTGCTCTTCACTTTC  
 ACTAACAGTAGCAACGGTCCGAACCTCATAACAACACTCAAACAAATTCTCAAGCGCTTCA  
 CAACCAATTGCCCTCTAACGTTCATGATAACTCATGAATAATGAAATCAGGCTAGT  
 AAAATTGATGATGGAATAATTCAAAACACTGTACCTGGTGGACGCTCCCTATACTGAGTC  
 TATAACCGCTTGGAAACTACAGGGATGTTAATACCAACTACAATGGATGATGATAT  
 AACTATCTATTGATGATGAGATAACCCACAAACCAAAAAAGAGGGTGGGTCGATC  
 ACAAGTTGTAACAAAAGCAGGCTTGTGACCCCGGGAAATTGAGATCTACTAGTGC  
 CGCACCGTACCCAGCTTCTGTACAAAGTGGTACGTCAGCTCCCTATACTGAGTC  
 TATTACACTGGCGCTGTTACAACGTCGTGACTGGAAAACACCGGTGAGCTCTAAGT  
 AAGTAACGGCCGCCACCGCGGTGGAGCTTGGACTTCTCGCCAGAGGTTGGTCAAGTC  
 TCCAATCAAGGGTGTGGCTTGTACCTTGCCAGAAATTACGAAAGATGGAAAAGGG  
 TCAAATCGTTGGTAGATACTGTTGACACTCTAAATAAGCGAATTCTTATGATTAT  
 GATTTTATTAAATAAGTATAAAAAAAATAAGTATACAAATTAAAGTGA  
 CTAGGTTTAAACGAAAATTCTGTTCTGAGTAACCTTCTCTGAGGTTGCT  
 TTCTCAGGTATAGCATGAGGTCGCTTATTGACCACACCTACCGGCATGCCAGCAA  
 ATGCCCTGAAATCGCTCCCCATTACCCAATTGAGATATGCTAACTCCAGCAATGAGT  
 TGATGAAATCTCGGTGTATTTATGTCCTCAGAGGACAATACCTGTTGTAATCGTTCT  
 CCACACGGATCCGATCAGGCGAAATTGTAACACGTTAATATTGTTAAATTCGCGTTA  
 AATATTGTTAAATCAGCTCATTTTAAACCAATAGGCCGAAATCGGAAACATCCCTTAT  
 AAATCAAAGAATAGACCGAGATAAGGGTTGAGTGTGTTCCAGTTGGAAACAAGAGTCCA  
 CTATTAAAGAACGTGGACTCCAACGTCAAAGGGCAAAACCGTCTATCAGGGCGATGGC-

FIGURE 98C

236 / 240

CCACTACGTGAACCATCACCTAATCAAGTTTTGGGGTCGAGGTGCCGTAAAGCACTA  
AATCGGAACCTAAAGGGAGCCCCGATTAGAGCTTGACGGGAAAGCCGGCGAACGTG  
GCGAGAAAGGAAGGGAAAGAAAGCGAAAGGAGCAGCGCTAGGGCGCTGGCAAGTGTAGCG  
GTCACGCTGCGCTAACCAACACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCC  
CATTCGCCATTCACTGCA

FIGURE 98D

237/240

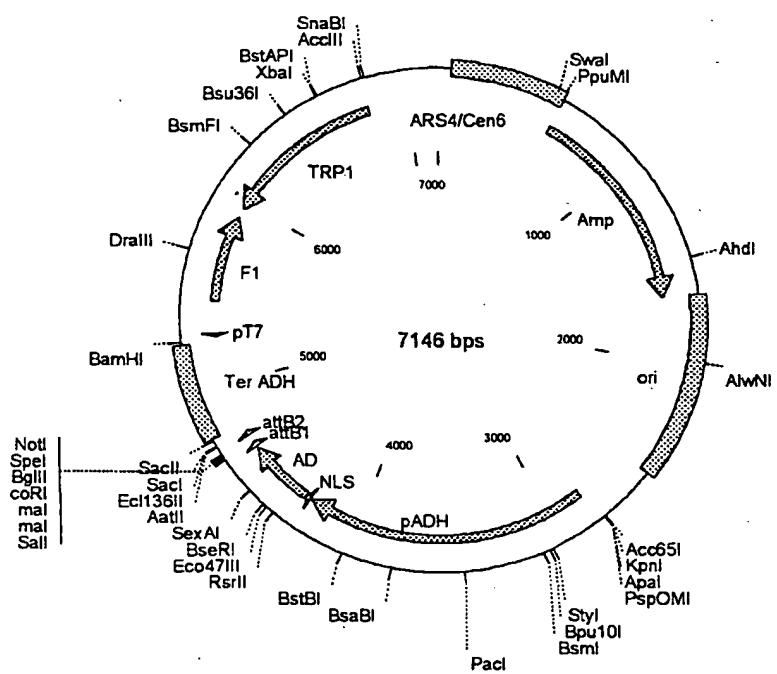
**pMAB86**

FIGURE 99A

238/240

pMAB86 7146 bp

GACGAAAGGGCCTCGTATA CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTT  
 CTTAGGACGGATCGCTGCCGTAACTTACACGCCCTCGTATCTTTAATGATGGAATA  
 ATTTGGGAATTACTCTGTGTTATTATTTATGTTGATTTGGATTTAGAAAGT  
 AAATAAAGAAGGTAGAAGACTACGGAATGAAGAAAAAAAATAAACAAAGGTTAAAAA  
 ATTTCAACAAAAGCTACTTACATATATTTATTAGACAAGAAAGCAGATTAAATA  
 GATATACATTGATTAACGATAAGTAAATGAAACATCACAGGATTTCTGTGTTGGTCT  
 TCTACACAGACAAGATGAAACAACTCCGCATTAATACCTGAGAGCAGGAAGAGCAAGATA  
 AAAGGTAGTATTGTTGGCGATCCCCCTAGAGTCTTTACATCTCCGAAAACAAAAACT  
 ATTTTTCTTAATTCTTTTACTTCTATTTTAATTATTTATTTATTTATTTAAAAA  
 ATTTAAATTATAATTATTTTATAGCAGTGTGAAAGGACCCAGGTGGCATTTCGG  
 GGAAATGTGCGCGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCCG  
 CTCATGAGACAATAACCCCTGATAAAATGCTCAATAATTGAAAAGGAAGAGTATGAGT  
 ATTCAACATTCCGTGTCGCCCTTATTCCCTTTTGCGGCATTTCGCTTCTGTGTTT  
 GCTCACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGATCAGTGGTGCAAGACTG  
 GGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTGAGAGTTTCGCCCCGAAGAA  
 CGTTTCCAATGATGAGCACTTTAAAGTCTGCTATGTGGCGGGTATTATCCGTATT  
 GACGCCGGCAAGAGCAACTCGGTGCCCATACACTATTCTCAGAATGACTGGTTGAG  
 TACTCACCAAGTCACAGAAAAGCATCTACGGATGGCATGACAGTAAGAGAATTATGCAGT  
 GCTGCCATAACCAGTGAATACACTGCGGCCACTTACTCTGACAAACGATCGGAGGA  
 CGAAGGAGCTAACCGCTTTTCAACATGGGGGATCATGTAACCTCGCCTGATCGT  
 TGGGAAACGGAGCTGAATGAAGCCATACCAAAAGACGAGCGTGACACCACGATGCCGTGA  
 GCAATGGCAACAACTGGTGCCTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCGG  
 CAACAATTAAAGACTGGATGGAGGCGATAAAAGTTGAGGACCCTCTCGCCTCGGCC  
 CTTCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGT  
 ATCATTGCACTGGGCCAGATGTTAAGCCCTCCGTATGTTAGTTATCTACACGACG  
 GGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCACTG  
 ATTAAGCATTGGTAACTGTCAGACCAAGTTACTCATATATACTTTAGATTGATTAAAAA  
 CTTCATTTTAATTAAAAGGATCTAGGTGAAGATCCTTTGATAATCTCATGACCAAA  
 ATCCCTTAACGTGAGTTTCTGTTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGA  
 TCTTCTGAGATCCTTTTCTGCGCGTAATCTGCTGCTGCAAACAAAAAACACCG  
 CTACCAAGCGGTGGTTGTTGCCGGATCAAGAGCTACCAACTCTTCCGAAGGTAAC  
 GGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTCTAGTGTAGCGTAGTTAGGCCAC  
 CACTTCAAGAAACTCTGTAGCACCCTACATACCTCGCTGCTGTAATCCTGTTACCAGTG  
 GCTGCTGCCAGTGGCGATAAGTCGTGCTTACGGGTTGGACTCAAGACGATAGTTACCG  
 GATAAGCGCAGCGTGGGTGAACGGGGGGTTCGTGACACAGCCCAGCTGGAGCGA  
 ACGACCTACACCGAAGTGAAGATACCTACAGCGTGGAGCATTGAGAAAGCGCACGCTTCCC  
 GAAGGGAGAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTGGAAACAGGAGAGCGCAG  
 AGGGAGCTCCAGGGGGAACGCGCTGGTATCTTATAGTCTGTCGGGTTGCCACCTC  
 TGACTTGAGCGTCGATTTGTATGCTCGTCAGGGGGCCGAGCCTATGAAAAACGCC  
 AGCAACCGGGCTTTTACGGTTCTGGCCTTGTGCTGGCTTTGCTCACATGTTCTT  
 CCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTGAGTGAGCTGATACC  
 GCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC  
 CCAATACGCAAACCGCCTCTCCCGCGCTGGCGATTATTAATGCACTGGCACGAC  
 AGGTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTATGTGAGTTACCTCACT  
 CATTAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCTATGTTGTGGAATTGTG  
 AGCGGATAACAATTACACAGAACAGCTATGACCATGATTACGCCAAGCTCGGAATT  
 AACCTCACTAAAGGAACAAAGCTGGTACCGGGCCCCCTCGAGATCCGGGATCGA  
 AGAAATGATGGTAAATGAAATAGGAATCAAGGAGCATGAAGGCAAAAGACAAATAAG  
 GTGCGAACGAAAATAAGTGAAGAAGTGTGATATGATGTTAGTGTGTTGGCGCCGA  
 AAAAACGAGTTACGCAATTGACAAATCATGCTGACTCTGTGGCGGACCGCGCTTGC  
 CGGCCCGCGATAACGCTGGCGTGAGGCTGTGCCGGAGTTTTGCGCCTGCATT  
 TTCCAAGGTTACCTGCGCTAAGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGG  
 GTTGCAGTGTGACGACCACGACAACGGTGTGATTATTAAGTGTGCGAAAGAACCTG  
 AGTGCATTGCAACATGAGTATACTAGAAGAATGAGCCAAGACTTGGCAGACCGCAGTT  
 GCCGGTGGTGCAGAACATAGAGCGACCATGACCTTGAGAGTGAAGGTGAGACCGCATAACCGCTA-

FIGURE 9B

239/260

GAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATA  
 CAACACTGGAAATGGTTGCTGTTGAGTACGCCTTCATTCAATTCTGGGTGTGCACTTTA  
 TTATGTTACAATATGGAAGGGAACTTTACACTTCTCCTATGCACATATAATTAAATTAAAGT  
 CCAATGCTAGTAGAGAAGGGGGTAACACCCCTCCGCGCTCTTCCGATTTTTCTAA  
 ACCGTGGAATATTCCGGATATCCTTGTGTTCCGGGTGACAATAATGGACTTCCCTCT  
 TTTCTGGCAACCAAACCCATACATCGGGATTCTTACATAATACCTCGTGGTCTCCCTAAC  
 ATGAGGTGGCGGAGGGAGATATAACATAGAACAGATACCAGACAAGACATAATGGGCT  
 AAACAAGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACATAACTG  
 TAGGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTCACTACCCCTTTCCATTGGC  
 ATCTATTGAAGTAATAATAGGCGCATGCAACTCTTTCTTTTTCTCTCTCT  
 CCCCGTGTGTCCTCACCATATCGCAATGACAAAAAAATGATGGAAGACACTAAAGGA  
 AAAAATAACGACAAAGACAGCACCAACAGATGCGTTCCAGAGCTGATGAGGGTA  
 TCTTCGAACACACGAAACTTTCTCTCCTCATTACCCACACTACTCTTAATGAGCA  
 ACGGTATACGGCCTCTTCCAGTTACTTGAAATTGAAAAAAAGTTGCCGCTTGC  
 CTATCAAGTATAAAATAGACCTGCAATTATACTTTGTTCTCGTCAATTGTTCTCGT  
 TCCCTTCTCCTGTTCTGCAAAATATTCAAGCTATACCAAGCATACAATC  
 AACTCCAAGCTTATGCCAAGAAGAAGGGAGGTCTCGAGCGGCCAATTAAATCAA  
 AGTGGGAATATTGCTGATAGCTATTGCTCTTCAACTAACAGTAGCAACGGTCCG  
 AACCTCATAAACAACCAAACATTCTCAAGCGCTTCAACACCAATTGCCCTCTAAC  
 GTTCATGATAACTTCATGAATAATGAAATCAGGCTAGTAAAATTGATGATGTAATAAT  
 TCAAAACACTGTCACCTGTTGGACGGACAAACTGCGTATAACGGTTGGAATCACT  
 ACAGGGATGTTAATACCACTACAATGGATGATGTTATAACTATCTATTGATGATGAA  
 GATACCCCACCAAACCAAAAAAGAGGGTGGTCATACAAGTTGTACAAAAAAAGCA  
 GGCTTGTGACCCCCGGGAATTCAAGATCTACTAGTGCACCGTACCCAGCTTCT  
 TGTACAAAGTGGTACGCTCAAGTAAGTAACGGCCACCGCGTGGAGCTTT  
 GGACTTCTCGCCAGAGGTTGGTCAAGTCTCAATCAAGGTTGTCGGCTTGTCTACCTT  
 GCCAGAAATTACGAAAAGATGAAAAGGGTCAAATCGTTGGTAGATACTGTTGACAC  
 TTCTAAATAAGCGAATTCTTATGATTATGATTTTATTAAATAAGTTAAATTTAA  
 AATAAGTGTATAAAATTAAAGTACTCTAGGTTAAAACGAAAATTCTGTTCT  
 GAGTAACCTTTCTGTAGGTCAAGGTTCTCAGGTATAGCATGAGGTGCTCTTAT  
 TGACCCACACCTCTACCGGATGCCGAGCAAATGCCCTGAAATCGCTCCCCATTCA  
 ATTGTAGATATGTAACCTCAGCAATGAGTTGATGAATCTCGGTGTGTTATGTC  
 CAGAGGACAATAACCTGTTGAATCGTTCTCACCGATCCAATTGCCCTATAGTGA  
 GTCGTATTACAATTCACTGGCGTGTGACTGGAAAACCCCTGGCGT  
 TACCCAACTTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGA  
 GGCCCGCACCGATGCCCTCCAAACAGTTGCCAGCCTGAATGGGAATGGACGCGCCC  
 TGTAGCGCGCATTAAGCGCGGGTGTGGTGGTACGCCAGCGTACCGCTACACTT  
 GCCAGCGCCCTAGCGCCGCTCTTCGCTTCTCCCTTCTGCCACGTTGCC  
 GGCTTCCCCGTCAGCTCTAAATCGGGGCTCCCTTAGGGTCCGATTTAGTGCTTA  
 CGGCACCTCGACCCAAAAAACTTGATTAGGGTGTGGTACGTAAGTGGCCATGCC  
 TGATAGACGGTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTG  
 TTCCAAACTGGAACAACACTCAACCCATCTCGGTCTATTCTTGTGTTATAAGGGATT  
 TTGCGGATTTCGGCTATTGGTTAAAAAATGAGCTGATTAAACAAAATTAAACGCGAAT  
 TTTAACAAAATTAAACGTTACAATTCTGATGCCGTGTTCTCCTTACGCATCTGT  
 GCGGTATTTCACACCGCAGGCAAGTGCACAAACAAACTTAAATAACTCAGTAA  
 TAACCTATTCTAGCATTGACGAAATTGCTATTGTTAGAGTCTTTACACCAT  
 TTGCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTAACTAAGCGCATT  
 CAACATTCTGCGCTAGTCCACCGCTAACATAAAATGTAAGCTTCCGGGCTCT  
 GCCTCCAACCCAGTCAGAAATCGAGTTCAACCCAAAGTTCACCGTCCACCTGCTT  
 CTGAATCAAACAAGGGATAAACGAATGAGGTTCTGTGAAGCTGCACTGAGTAGTATGT  
 TGCAGTCTTGTGAAATACGAGTCTTTAATAACTGGCAAACCGAGGAACCTTGGTATT  
 CTTGCCACGACTCATCTCATGCCAGTTGGACGATATCAATGCCGTAACTATTGACCAGAG  
 CCAAAACATCCTCTTAGGTTGATTACGAAACACGCCAACAGTATTGAGTGCCTG  
 AACTATTGTTATATGTTTACAAGACTTGAAATTTCCTGCAATAACCGGGTCAATTG  
 TTCTCTTCTATTGGGCACACATATAACCCAGCAAGTCAGCATCGGAATCTAGAGCAC  
 ATTCTGCGGCCTCTGTGCTGCAAGCCGAAACTTCACCAATGACCAGAAACTACCTG  
 TGAAATTAAATAACAGACATACTCCAAGCTGCCATTGTGCTTAATCACGTATACTC  
 TGCTCAATAGTACCAATGCCCTCCCTTGGCCCTCTCCTTTCTTGTGACCGAAT-

FIGURE 99C

240 / 240

TAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGTACGTAAGGTGACAAG  
CTATTTTCATAAAGAAATATCTTCCACTACTGCCATCTGGCGTCATAACTGCAAAGTAC  
ACATATATTACGATGCTGTCTATTAAATGCTTCTATATTATATATAGTAATGTCGTT  
TATGGTGCACTCTCAGTACAATCTGCTTGATGCCGCATAGTTAACGCCAGCCCCGACACC  
CGCCAACACCCGCTGACGCCCTGACGGGCTTGTCTGCTCCGGCATCCGCTACAGAC  
AAGCTGTGACCGTCTCCGGGAGCTGCATGTTCAGAGGTTTCACCGTCATCACCGAAAC  
GCGCGA

FIGURE 99D

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

REC'D

A. The indications made below relate to the microorganism referred to in the description on page 54, line 8.

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution

Agricultural Research Culture Collection (NRRL)  
 International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
 Peoria, Illinois 61604  
 United States of America

Date of deposit February 27, 1999

Accession Number

NRRL B-30103

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet 

Escherichia coli DB3.1(pEYC15101)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

 This sheet was received with the international application
 
 This sheet was received by the International Bureau on:
 

Authorized officer



Authorized officer

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)  
 International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
 Peoria, Illinois 61604  
 United States of America

Date of deposit February 27, 1999

Accession Number

NRRL B-30100

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)

This information is continued on an additional sheet

Escherichia coli DB3.1(pENTR-1A)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

This sheet was received with the international application

This sheet was received by the International Bureau on:

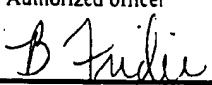
Authorized officer



Authorized officer

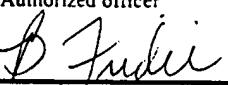
**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>55</u>, line <u>16</u></p>		
<p><b>B. IDENTIFICATION OF DEPOSIT</b></p>		<input checked="" type="checkbox"/> Further deposits are identified on an additional sheet
<p>Name of depositary institution  Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>)  1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
		NRRL B-30102
<p><b>C. ADDITIONAL INDICATIONS</b> (<i>leave blank if not applicable</i>)</p>		<input type="checkbox"/> This information is continued on an additional sheet
<p>Escherichia coli DB3.1(pENTR-3C)</p> <p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p><b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (<i>if the indications are not for all designated States</i>)</p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS</p> <p><i>(leave blank if not applicable)</i></p> <p>The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer 		Authorized officer	

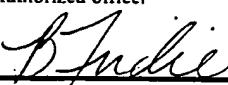
**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>55</u>, line <u>16</u>.</p>		
<p><b>B. IDENTIFICATION OF DEPOSIT</b></p>		<input checked="" type="checkbox"/> Further deposits are identified on an additional sheet
<p>Name of depositary institution  Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>)  1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
		NRRL B-30101
<p><b>C. ADDITIONAL INDICATIONS</b> (<i>leave blank if not applicable</i>)      This information is continued on an additional sheet <input type="checkbox"/></p>		
<p>Escherichia coli DB3.1(pENTR-2B)</p> <p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p><b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (<i>if the indications are not for all designated States</i>)</p>		
<p><b>E. SEPARATE FURNISHING OF INDICATIONS</b> (<i>leave blank if not applicable</i>)</p> <p>The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer </p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
--	--

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

A. The indications made below relate to the microorganism referred to in the description on page <u>WPO</u> <u>RCT</u> <u>20-21</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input checked="" type="checkbox"/></span>	
Name of depositary institution  Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution ( <i>including postal code and country</i> )  1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit	February 27, 1999
Accession Number	NRRL B-30108
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
Escherichia coli DB10B(pCMV Sport6)	
In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )	
The indications listed below will be submitted to the international Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

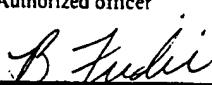
**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>54</u>, line <u>9</u>.</p>		
<p><b>B. IDENTIFICATION OF DEPOSIT</b></p>		<input checked="" type="checkbox"/> Further deposits are identified on an additional sheet
<p>Name of depositary institution  Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>)  1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
NRRL B-30105		
<p><b>C. ADDITIONAL INDICATIONS</b> (<i>leave blank if not applicable</i>)</p>		<input type="checkbox"/> This information is continued on an additional sheet
<p>Escherichia coli DB3.1(pEZA15103)</p>		
<p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p><b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (<i>if the indications are not for all designated States</i>)</p>		
<p><b>E. SEPARATE FURNISHING OF INDICATIONS</b> (<i>leave blank if not applicable</i>)</p>		
<p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer 		Authorized officer	

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**  
**(PCT Rule 13bis)**

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>54</u>, line <u>9</u>.</p>		
<p><b>B. IDENTIFICATION OF DEPOSIT</b></p>		<input checked="" type="checkbox"/> Further deposits are identified on an additional sheet
<p>Name of depositary institution  Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>)  1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
		NRRL B-30104.
<p><b>C. ADDITIONAL INDICATIONS</b> (<i>leave blank if not applicable</i>)</p>		<input type="checkbox"/> This information is continued on an additional sheet
<p>Escherichia coli DB3.1(pEZA15102)</p>		
<p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p><b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (<i>if the indications are not for all designated States</i>)</p>		
<p><b>E. SEPARATE FURNISHING OF INDICATIONS</b> (<i>leave blank if not applicable</i>)</p>		
<p>The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

<p>For receiving Office use only</p>		<p>For International Bureau use only</p>	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
<p>Authorized officer</p> 		<p>Authorized officer</p>	

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM  
(PCT Rule 13bis)**

A. The indications made below relate to the microorganism referred to in the description on page 52, line 31

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)  
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street  
Peoria, Illinois 61604  
United States of America



Date of deposit	February 27, 1999	Accession Number	NRRL B-30099
-----------------	-------------------	------------------	--------------

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

This sheet was received with the international application

This sheet was received by the International Bureau on:

Authorized officer Barbara Fridie  
PCT Operations - IED Team 1  
703-305-3737 (T) 703-305-3230 (FAX)

Authorized officer

*Escherichia coli DB3.1(pENTR-3C)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-3C)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-2B)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pENTR-2B)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-2B)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-1A)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pENTR-1A)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-1A)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSPORT6)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSPORT6)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSport6)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15103)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15103)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15103)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15102)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15102)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZA15102)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15101)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15101)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

**NORWAY**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

**SINGAPORE**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15101)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-3C)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**CANADA**

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/05432

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.  
US CL :435/912, 252.3, 320.1; 530/350; 536/ 23.1, 24.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/912, 252.3, 320.1; 530/350; 536/ 23.1, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ---	US 5,888,732 A (HARTLEY et al.) 30 March 1999, see entire document.	1-21, 25-30 36-38 -----
Y,P		22-24, 31-35
X - Y	HASAN et al. Escherichia coli genome targeting, I. Cre-lox-mediated in vitro generation of ori- plasmids and their in vivo chromosomal integration and retrieval. Gene. 1994, Vol. 150, pages 51-56, see entire document.	1-5, 10, 11, 19-21 ----- 15-18, 22-38
X - Y	KATZ et al. Site-specific recombination in Escherichia coli between the att sites of plasmid pSE211 from Saccharopolyspora erythraea. Mol. Gen. Genet. 1991, Vol. 227, pages 155-159, see entire document.	1-11, 19-21 ----- 15-18, 22-38

Further documents are listed in the continuation of Box C.  See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
*O* document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
08 MAY 2000	23 MAY 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer <i>Christina Lawrence Fox</i> IREM YUCEL
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/US00/05432

**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ASTUMIAN et al. Site-specific recombination between cloned attP and attB sites from the Haemophilus influenzae bacteriophage HP1 propagated in recombination deficient Escherichia coli. J of Bacteriology. March 1989, Vol. 171, No. 3, pages 1747-1750, see entire document.	1-11, 19-21 -----
-		
Y		15-18, 22-38

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/05432

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C07K 1/00, 14/00; C12N 1/21, 15/00, 15/09, 15/63, 15/70; C12P 19/34

## B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, STN (CAPLUS); DIALOG (MEDLINE, BIOSIS, SCISEARCH, PASCAL)

Terms: att (B?, P?, R?, L?), MCS, POLYLINKER, PLASMID, VECTOR, LOCALIZATION, SIGNAL, TRANSCRIPTION, TERMIN?, TRANSLATION?, ORI, REPLICON, GST, HEXHIST?, THIOREDOX?, CLEAVAGE, SITE?, SPECIF?, DIRECT?, RECOMBIN?, CLON?, INSERT?

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**